



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2017

Mind and rights: Neuroscience, philosophy, and the foundations of legal justice

Mahlmann, Matthias

DOI: <https://doi.org/10.1017/9781108355223.006>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-144880>

Book Section

Originally published at:

Mahlmann, Matthias (2017). Mind and rights: Neuroscience, philosophy, and the foundations of legal justice. In: Sellers, M N S. Law, Reason, and Emotion. New York: Cambridge University Press, 80-137.

DOI: <https://doi.org/10.1017/9781108355223.006>

Mind and Rights: Neuroscience, Philosophy, and the Foundations of Legal Justice

Matthias Mahlmann

I REASON, CONSCIENCE, AND RIGHTS¹

The Universal Declaration of Human Rights (UDHR) famously begins with an anthropological assumption: All human beings, it asserts, “are endowed with reason and conscience.”² There are no explicit conclusions drawn from this in its text, but it is clear that these assumed properties are taken to be relevant to the idea of human rights that the Universal Declaration restates powerfully after the many disasters of the Second World War. It is not very far-fetched to think that one underlying idea is that *because* of these properties of human beings, alone or in conjunction with others, it is justified to conclude that humans are endowed with certain inalienable rights.³ Another possible reading is that because of “reason and conscience” humans are in fact in the epistemic position to *understand* that their human rights are justified, that they are not doomed to ignorance and consequently should take action to protect them.⁴

¹ Preliminary versions of parts of the materials have been presented at the Law and Neuroscience Conference, December 2014 in Zurich and the Legal Philosophy Colloquium of the Cluster of Excellence Normative Orders in Frankfurt/M and at the University of Utrecht. I would like to thank in particular Noam Chomsky, Marcus Düwell, Klaus Günther, Lutz Jäncke, Christopher McCrudden, John Mikhail, Hubert Rottleuthner, Lutz Wingert, and Eyal Zamir for discussion of these issues and critical (and very careful) comments.

² Art. 1, UN General Assembly, Universal Declaration of Human Rights, 10 December 1948, 217 A (III).

³ Cf. the account of the debate on these terms in the *travaux préparatoires*, where among other things the question was discussed whether the reference to reason and conscience would exclude certain human beings, J. Morsink, *The Universal Declaration of Human Rights*, 1999, p. 296ff. The point of these terms was to identify distinguishing properties of human beings; cf. *ibid.*

⁴ The preamble of the UDHR states, after all, that “disregard and contempt for human rights have resulted in barbarous acts which have outraged the *conscience of mankind*” (emphasis by

These two thoughts may appear as a truism to the average observer and of no particular concern. What else, one may be tempted to say, was to be stated after half of the world and the ethical foundations of an epoch were in still-smoking ruins? Human beings are beastly predators driven by the irrational will to power? Protean beings,⁵ “nicht festgestellte Tiere,” animals without a fixed nature,⁶ that are the mere playthings of historical and social change? Shrewd calculators of interests driven by the sole motive of maximizing their own idiosyncratic preferences? Would this not have missed the lessons and demands of a pivotal historic moment?

That these thoughts about the justification of human rights and the epistemic condition of human beings are in fact more than lofty rhetoric is, however, far from clear if one considers contemporary debates about reason and rights. What the foundations of rights are is as much debated as the claim that there are universal rights of every human everywhere, a conclusion that the Universal Declaration draws as one of its defining elements. Reason has become a notoriously contentious concept in many quarters of reflection, and conscience as a human property of constitutive importance for moral orientation does not necessarily fare better.

Against this background, many problems exist that are worthy of the serious attention and the admirable work devoted to them in contemporary human rights theory. Given the anthropological assertion in the Universal Declaration, the following question may also come to mind: What is the actual relationship between human thought, its structure and exercise, and the idea of human rights, which is surely among the most important products of human thinking? This question will be explored below.

To this end, some more thoughts have to be devoted to the question of why the relationship between the human mind and rights is of some theoretical interest. So, the first question will be: *Why does the theory of mind matter for ethics and law?*⁷ Second, the concept or idea of a human right as a subclass of moral and legal subjective (or in a different terminology – claim) rights used

the author). For a predominantly epistemological reading J. Morsink, *The Universal Declaration of Human Rights*, 1999, p. 296ff.

⁵ Cf. R. Rorty, Human Rights, Rationality and Sentimentality, in: S. Shute/S. Hurley (eds.), *On Human Rights*, 1993, p. 111ff, 115.

⁶ F. Nietzsche, *Jenseits von Gut und Böse*, Aphorismus 62, in: G. Colli/M. Montinari (eds.), *Sämtliche Werke*, Bd. 5, 1999.

⁷ The term “ethics” is used in a variety of ways. One is to take ethics as a reflective theory of morality. Another widespread understanding of the term is to use “ethics” for anything that concerns the good, flourishing life in a roughly eudaemonistic sense, e.g., J. Habermas, *Faktizität und Geltung*, 1992, p. 139ff; R. Dworkin, *Justice for Hedgehogs*, 2011, p. 13ff. Here, “ethics” is used in the former, not the latter sense.

will be outlined and clarified. This analysis of the concept or idea of rights is indispensable to answer the question: *What precisely are we talking about?* Third, the question *Where do rights come from?* will occupy attention just long enough to substantially understand why a reflection on the historical genealogy of human rights leads necessarily beyond the limits of human rights history into the deep waters of the epistemology and ontology of human rights and thus to those kinds of problems these remarks intend to explore. History or historicism, it is argued, offers no escape from them. To this end, one of the two currently particularly interesting fundamental forms of human rights' revisionism, the historical, genealogical attack on human rights, will be used as a topical and helpful expository device. Fourth, the question *Why are rights justified?* will be considered. The assumption behind this discussion is that there is no meaningful epistemology of human rights without a normative theory of how they can be justified. This is because the latter formulates the claims the epistemological merits of which are to be assessed by the former. Fifth, after having sufficiently prepared the ground by the preceding remarks, the core issue of these reflections can be addressed: *What is, after all, the importance of the theory of mind for the project of human rights?* Here the second fundamental challenge to the idea of human rights will be discussed. This attack stems from today's neuroscientific neo-emotivism, which is interesting in itself and has the advantage that the critique of this form of human rights revisionism has considerable heuristic merits for a constructive account of the theory of mind and the foundations of human rights. How a theory of human rights could draw from the theory of mind, and more concretely from a mentalist account of ethics and law, to provide such a constructive account is the final perspective to be explored.

There are very serious contemporary political, cultural, and theoretical reasons to worry about the project of human rights. One should not take the existence of the level of civilization epitomized by human rights for granted. The history of the last century is sobering. Massive crimes were committed because fantastic ideologies like National Socialism held their barbarous sway. Camus called it with good reasons "le siècle de la peur",⁸ a century of fear that formulates the categorical imperative: "ni victimes, ni bourreaux," neither to become victims, nor hangmen.⁹ Given this and the added experience of the many ways of suffering after this cataclysm around the world, the recent past has certainly taught the lesson not to put too much confidence in the decent

⁸ A. Camus, *Combat*, 19 November 1946, in: *Cahiers Albert Camus* 8, Camus à Combat, ed. J. Lévi-Valensi, 2002, p. 608.

⁹ Ibid.

behavior of human beings. To be sure, such skepticism does not necessarily imply a verdict about reason, or an endorsement of theories of its intrinsic dark side¹⁰ or of the amoral driving forces of the human will.¹¹ But it does nourish the very ancient reluctance to underestimate the fragility of civilization. The Athenians did not lack culture but still sowed destruction in the Peloponnesian Wars, for others and ultimately for themselves.¹²

It is thus not only a theoretically important but also practically crucial end to dispel doubts about the justification of human rights. It is necessary to strengthen the not at all self-evident motivation to do something to defend their fragile rule where it exists and to help to increase their sway, which challenges power, injustice, and bondage in this world. Whether a reflection on mind and rights can contribute to these large questions is asked in the five steps of this reflection. The first is to better clarify why the enquiry into mind and rights matters for the theory of human rights.

II HUMAN RIGHTS AND HUMAN THOUGHT – WHY DOES IT MATTER?

The topic of mind and rights will not necessarily strike everybody as likely to be productive. This topic may seem untimely or to be of only obscure theoretical merit for the contemporary understanding of human rights, given the troubled times we live in, full of violent conflict on a massive scale, and deep social problems. The war in Syria, to take an obvious example, will devastate the country for generations to come. New political monstrosities such as ISIS have been borne and nurtured in violent religious sectarianism. The brooding threat of violent attacks against civilian targets prolongs the era of 9/11 into an uncanny future. In Ukraine, a civil war conjures up the specter of cruel proxy wars of the Cold War area. In the aftermath of the Arab Spring, a state like Libya has collapsed, and others like Egypt seem to be drifting back into an authoritarian future that was their past. Mass migration is turning the Mediterranean into a graveyard. Conflicts in Nigeria with Boko Haram or in the Congo gain less attention, though they merit as much consideration as any human tragedy

¹⁰ Following, e.g., the argument of M. Horkheimer/T. W. Adorno, *Die Dialektik der Aufklärung*, 1969.

¹¹ Cf. with the consequence to deny the “Wille zum Leben,” will to life, A. Schopenhauer, *Die Welt als Wille und Vorstellung*, Sämtliche Werke, ed. W.v. Löhneysen, 1986, Bd. 1, § 68; Bd. 2, Chapter 48.

¹² A perceptive commentator like Thucydides took it as a premise and justification of his work that humans will repeat such miseries as the war he describes and may therefore profit from a true account of the past, cf. idem., *History of the Peloponnesian War*, 1928, Book 1, XXII.

anywhere on this globe. An unresolved international economic crisis may cause much upheaval and further injustice in many parts of the world. And problems like global warming stay unresolved despite robust insights into the causes and the steps that should be taken to prevent environmental disasters, which will on all accounts have a major impact on human rights in the future.

And there is another possible concern about the timeliness of the question pursued. Human rights have been under political attack for years. That is in principle nothing new. Fundamental rights play many roles in human history, but not least among these is the threat that they pose as an ethical idea and legal instrument subversive to political power. Fundamental rights are, in addition, not only unpleasant for any aspiration to unfettered might, because they draw limits to its exercise; they are unpleasant for others, too, not least for – often comfortably self-righteous – social majorities. Human rights draw lines that limit the capability of powerful people to impose their view of the right and proper on other people as well. Human rights are therefore a precious asset for any minority, the dissidents and outsiders, or simply the weakest parts of human associations. What is unpleasant for political power and social majorities will give rise to enemies, and human rights have plenty of them.

There are in addition to these familiar enemies of human rights particular new threats from within democratic and constitutional systems. Prime examples of these are the measures taken in the so-called war on terror, the erosion of such fundamental norms as the prohibition of torture, the profound threat to privacy and self-determination through the international surveillance systems that came to light in the wake of the disclosure of NSA documents in which many countries and not only the “Five Eyes” participated; the undermining of an international rule of law through ongoing practices of extrajudicial killings through drone strikes¹³ or the symptoms of a new religious illiberalism epitomized by the Swiss ban on the minaret.

Given this state of affairs, a reflection on human rights needs no excuse. But why mind and rights? Is there nothing more important to think about, when one wants to talk about human rights, than such an airy topic? Given just these examples – are the problems of mind and rights not just a dubious pastime for the idle lodgers in the ivory tower? Would an enquiry into the politics of human rights; their role in violent conflicts and proxy wars like the ones

¹³ On these latter two examples and their wider impact cf., e.g., D. Cole, Must Counterterrorism Cancel Democracy?, *The New York Review of Books* 1 (2015), p. 26ff. On targeted killings cf. Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Philip Alston, Addendum, Study on targeted killings, United Nations, General Assembly, 28 May 2010, A/HRC/14/24/Add. 6.

mentioned above; their implications for the social dimensions of the economic crisis and the equity of solutions pursued for it; the difficult relation of human rights and religion or their function in the current display of international power (among so many other problems that come to mind) not be more appropriate approaches?

And are not other perspectives more promising if one wants to engage in the theory of fundamental rights, say rights and culture, rights and the narratives of modernity, the social construction of rights or the analysis of the social assemblage of rights? Is the question of the relationship between mind and rights thus not practically irrelevant and theoretically off the mark?

There are reasons to believe that this impression may be misleading and that it is on the contrary crucial for the understanding of human rights to tackle this problem. First of all, given the recent upsurge of interest in the relevance of neuroscience and empirical psychology for the understanding of ethics and law,¹⁴ it is not far-fetched at all, but rather quite necessary to consider in some detail the lessons that the modern theory of the human mind may have for the understanding of the foundations of human rights and – given the constitutive role of human rights for legal systems in general – for the foundations of legal justice. Neuroscience and psychology have a major impact on the contemporary reflection of normative issues. There is consequently no plausible theory of human rights imaginable without a thorough reflection on whether the contemporary understanding of the working of the human mind is or is not important for the understanding of human rights.

Secondly, the two kinds of human rights revisionism mentioned underline this point. Many arguments have been put forward against the idea of human rights, from Bentham's doubts¹⁵ to the attack on abstract and general norms under the theoretical auspices of a negative dialectic¹⁶ or postmodernity.¹⁷ Two recent challenges add critical perspectives to this line of thought.

The first is what one may call a genealogical revisionism with deconstructive normative connotations.¹⁸ This challenge formulates the thesis that

¹⁴ Cf., e.g., A. Appiah, *Experiments in Ethics*, 2008; M. Gazzaniga, *The Ethical Brain*, 2005; S. J. Morse/A. D. Roskies (eds.), *A Primer on Criminal Law and Neuroscience*, 2013; O. D. Jones, *Seven Ways Neuroscience Aids Law*, *Vanderbilt University Law School, Public Law and Legal Theory, Working Paper* 13 – 28 (2013); M. S. Pardo/D. Patterson, *Mind, Brain and Law*, 2013; or the casebook, O. D. Jones/J. D. Schall/F. X. Shen, *Law and Neuroscience*, 2014.

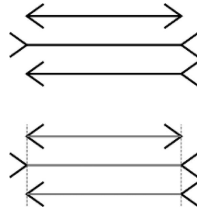
¹⁵ J. Bentham, *Anarchical Fallacies*, in: J. Waldron (ed.), *Nonsense upon Stilts*, 1987, p. 53 on natural rights.

¹⁶ T. W. Adorno, *Negative Dialektik*, 1997, p. 281.

¹⁷ J. Derrida, *Force of Law: The Mystical Foundations of Authority*, in: D. Cornell/M. Rosenfeld/D. Carlson (eds.), *Deconstruction and the Possibility of Justice*, 1992, p. 3ff, 13ff, 59ff.

¹⁸ S. Moyn, *The Last Utopia*, 2010.

human rights are of recent, historically contingent and politically doubtful origin that contaminates their legitimacy. The second is a neuroscientific, psychological revisionism that mobilizes what is perhaps the most influential model of the human mind – the dual process model of the mind¹⁹ – against the idea of human rights by arguing that this idea is just a product of a particular mechanism of the mind that must not, however, direct our decision-making. The consequence of such theories is far-reaching. Human rights appear as something like a cognitive illusion in the technical sense, i.e., a mental phenomenon that is necessarily experienced by human beings, given a certain input, because of the structure of the human mind, but that in fact delivers erroneous information about the true state of the world.²⁰ Beliefs about the validity of human rights are as necessarily experienced and as illusory as the impression when seeing the Müller-Lyer illusion:



Of the parallel lines, with fins pointing in different directions, the upper line appears to be shorter than the others, though in fact all are of equal length. This is a very well-known visual illusion, and despite everybody's knowing the effect, one still perceives it in this manner. In the same way, the idea of human rights is a cognitive illusion, it is argued: They are an offshoot of mental mechanisms that deliver certain ideas whether we want it to or not and whose erroneous nature can be rectified only by other forms of rational thinking. The impression of the justification of human rights, however, stays with us like the impression of the different lengths of the lines. We can thus not liberate ourselves entirely from such ideas as human rights, but we can learn to ignore them when it is important to do so.

These two critiques are helpful because they lead to two questions any theory of human rights has to answer. At first sight, the first kind of critique

¹⁹ Cf. D. Kahneman, *Thinking, Fast and Slow*, 2011.

²⁰ J. Greene, *Moral Tribes*, 2013. For a related but not identical argument C. Sunstein, *Moral Heuristics*, *Behavioural and Brain Sciences* 28 (2005), p. 531ff; and the comments by J. Mikhail, *Moral Heuristics or Moral Competence, Reflections on C. Sunstein*, *Behavioural and Brain Sciences* 28 (2005), p. 557ff.

concerns only matters of moral and legal history. On a deeper level, however, the legitimacy of human rights is at stake. If human rights have a partisan, politically suspect origin, this, according to this view, undermines their legitimacy. The implicit assumption is that the criteria for the validity of the claim that human rights are justifiable, and perhaps universally so, have a certain historical origin, though it is not formulated which origin would make human rights legitimate, only which origins render them illegitimate. This thesis can ultimately be criticized only by an alternative theory of the validity of claims about the justification of human rights, and to outline such a theory is impossible without an epistemology and ontology of human rights.

It is clear that the legitimacy of human rights cannot be justified without the assumption that the proposition “Human rights are (universally) valid” is in fact an act of cognition, an act of insight and not of error. This leads to the second challenge. This position argues that a proper state-of-the-art scientific understanding of the structure and cognitive conditions of the exercise of human rationality or reason does not yield arguments for the legitimacy of the idea of human rights, let alone for their universality. On the contrary, such theories assert that there are hard scientific grounds to believe in the necessity of the radical and irreverent critique of such ideas that are put without reason on the pedestal of admirable human achievements and properly belong in the dustbin of human thought. From such a perspective that takes fundamental principles of morality and the idea of human rights as something like a cognitive illusion, the specter of Descartes’ evil demon reappears.²¹ The cheating god he thought to be unimaginable now resides within us as part of our cognitive machinery. Our mind is not ultimately, despite all errors, an inner light, as Descartes among many hoped, that leads to insight.²² Quite the contrary, right in the heart of human understanding, sources of obscurantism cloud our vision, among them the idea of the morality of human rights. To assess the merits of such claims, one has to engage seriously with the question of the relationship between the theory of mind and human rights, and explore what one can learn from such theoretical attacks.

The fundamental question underlying these remarks is furthermore deeply embedded in a rich and fertile tradition of thought. It is a central Socratic assumption that is formative, not only for the so-called history of “Western” practical philosophy, that moral principles and especially justice are an object of true human insight.²³ In this reflection, the criteria for genuine insight in

²¹ R. Descartes, *Principia Philosophiae*, ed. A. Buchenau, 1922, I, 29, 30. ²² Ibid., I, 11, 30.

²³ Cf., e.g. Plato, *Gorgias*, 508e–509a. The disavowal of knowledge in this passage after the assertion of knowledge is best understood as “Socratic irony,” cf. G. Vlastos, *Socrates*, 1991, p. 21ff, 236ff.

contrast to unwarranted opinion or error were and are of central importance.²⁴ What is constitutive for justice and thus forms such a criterion is therefore a central question from pre-Socratic²⁵ thought to contemporary theories of justice.²⁶

Insight is, however, not just there. It has to be gained by humans through a mental process of reflection, of thinking, and thus through the active exercise of the instrument to gain insight: the human mind. A central question of modern epistemology is therefore whether the structure of the mind is important for what counts as criteria for insight. In other words: Is thought more than just a pure medium of cognition? Does it – by its properties and structure – determine what appears to humans as truth and not as error? Are the properties and the structure of the mind decisive for the content of thought? If that is so, these mental properties and structures have to be investigated to understand the nature and foundation of human cognition, including moral judgment and moral cognition. A reflection not only on the substance of insight but also on the mind that provides insight, however, is a classical project of modern philosophy. It is – despite different opinions about the structure of the mind – the common attempt of Descartes,²⁷ Locke,²⁸ and Leibniz²⁹ to describe the properties of human thought that enable human beings to understand, requiring “art and pains”, because “the understanding, like the eye, whilst it makes us see, and perceive all other things, takes no notice of itself.”³⁰ It is the Humean resolve to enquire into the “secret springs and principles, by which the human mind is

²⁴ Cf. Plato, *Euthyphron*, 5d, 6d – 6e.

²⁵ See, e.g., Simonides’ principles, Plato, *Politeia*, 331e.

²⁶ J. Rawls, *A Theory of Justice*, 1971; S. Gosepath, *Gleiche Gerechtigkeit*, 2004; A. Sen, *The Idea of Justice*, 2009.

²⁷ Cf., e.g., R. Descartes’ analysis of the human mind, *Principia philosophiae*, 1922, I, VIIIff.

²⁸ Cf. J. Locke, *An Essay Concerning Human Understanding*, ed. R. Woolhouse, 1997, Book I, Chapter I, § 1.

²⁹ G. W. F. Leibniz, *Nouveau Essais Sur l’Entendement Humain*, ed. W.v. Engelhardt/H.H. Holz, 1996, Preface: “Il s’agit de savoir, si l’Ame en elle-même est vuide entierement comme des Tablettes, où l’on n’a encor rien écrit (Tabula Rasa) suivant Aristote et l’Auteur de l’Essay, et si tout ce qui y est tracé vient uniquement des sens et de l’experience, ou si l’ame contient originairement les principes de plusieurs notions et doctrines que les objets externes reveillent seulement dans les occasions, comme je le croix avec Platon et même avec l’Ecole et avec tous ceux qui prennent dans cette signification le passage de S. Paul (Rom. 2, 15) où il marque que la loy de Dieu est écrite dans les cœurs.”

³⁰ J. Locke, *An Essay Concerning Human Understanding*, 1997, Book I, Chapter I, § 1: “The understanding, like the eye, whilst it makes us see, and perceive all other things, takes no notice of itself: and it requires art and pains to set it at a distance, and make it its own object. But whatever be the difficulties, that lie in the way of this inquiry; whatever it be, that keeps us so much in the dark to ourselves; sure I am, that all the light we can let in upon our minds; all the acquaintance we can make with our own understanding, will not only be very pleasant, but bring us great advantage, in directing our thoughts in the search of other things.”

actuated in its operations,”³¹ the Kantian project of a critique of reason, the demand of the “ripened power of judgment of our age, which will no longer be put off with illusionary knowledge, and which demands that reason should take on anew the most difficult of its tasks, namely, that of self-knowledge.”³² To be sure, this project has been the object of manifold critiques – from Heidegger’s doubts about the Cartesian metaphysics³³ to Wittgensteinian externalism.³⁴ But it could still turn out to have some merits to consider closely “the ways, whereby our understanding comes to attain those notions of things we have”³⁵ to avoid a philosophical enquiry beginning “at the wrong end.”³⁶

This approach has its parallel in the modern philosophy of language and its thesis that a theory of language is a necessary element of any theory of human knowledge. One of its persistent questions is whether the properties of ordinary languages influence the nature of human thought, making it relative to a certain language,³⁷ or on the contrary provide reasons that confirm the

³¹ D. Hume, *An Enquiry Concerning Human Understanding*, Sec. I, in: P. H. Nidditch (ed.), David Hume, *Enquiries*, 3rd ed., 1975, p. 14.

³² I. Kant, *Kritik der reinen Vernunft* (1st ed. 1781), Akademie Ausgabe, Bd. IV, p. 9: a demand “der gereiften Urteilskraft des Zeitalters, welches sich nicht länger durch Scheinwissen hinhalten läßt, und eine Aufforderung an die Vernunft, das beschwerlichste aller ihrer Geschäfte, nämlich das der Selbsterkenntnis aufs neue zu übernehmen,” translation: idem., *Critique of Pure Reason*, Cambridge Edition of the Works of Immanuel Kant, 1999, p. 100f.

³³ Cf. M. Heidegger, *Sein und Zeit*, 15th ed., 1979, p. 89ff on the deficient approach to the world of Cartesian metaphysics, especially missing its “Zuhandenheit” (being “ready-to-hand”).

³⁴ Cf. L. Wittgenstein, *Zettel*, ed. G.E.M. Anscombe/G.H.v. Wright, 2007, 606, 608; Putnam, *Pragmatism*, 1995, p. 79: “The mind is not in the head.” Concretely on the question of neuroscience and more generally for an externalist account of the mind, cf. M. R. Bennett/P. M. S. Hacker, *The Philosophical Foundations of Neuroscience*, 2003; M. S. Pardo/D. Patterson, *Minds, Brains, and the Law*, 2013, p. 12ff: rule following (with conceptual necessity) is not in the head.

³⁵ J. Locke, *An Essay Concerning Human Understanding*, 1997, Book I, Chapter I, § 2.

³⁶ Ibid., Book I, Chapter I, § 7: “For I thought that the first step towards satisfying several enquiries, the mind of man was very apt to run into, was, to take a survey of our own understandings, examine our powers, and see to what things they were adapted. Till that was done, I suspected we began at the wrong end, and in vain sought for satisfaction in a quiet and sure possession of truths, that most concerned us, while we let loose our thoughts into the vast ocean of being, as if all that boundless extent, were the natural and undoubted possession of our understandings, wherein there was nothing exempt from its decision, or that escaped its comprehension.”

³⁷ Cf. the classical thesis of W. v. Humboldt, *Über die Verschiedenheit des menschlichen Sprachbaus*, in: idem., *Werke*, ed. A. Flitner/K. Giel, 9th ed., 2002, Bd. III, p. 224 that language is an “eigenthümliche Weltansicht,” language is a kind of worldview. It is for him a central task of the study of language to determine the part that language plays in the creation of beliefs, ibid., p. 153. For the Sapir/Whorf hypothesis of the determination of thought by language cf. B. Whorf, *Language, Thought and Reality*, 1956, p. 212; on the implausibility of this thesis, e.g., S. Pinker, *The Language Instinct*, 1994, p. 59ff.

possibility of universally shared human thought.³⁸ Here too, not only the material criteria for insight but also the medium by which insight may be gained is of central importance and by some is even regarded as crucial for determining the reach and limit of human understanding.

III THE CONCEPT OF HUMAN RIGHTS

The concept or idea of a “right” or perhaps more precisely a “subjective right” is an intricate one. As will be illustrated in the historical reflection, it is important to distinguish words from what is designated by these terms. What is referred to by the term “right” or “subjective right” may be the meaning of many other terms, too. It can even be expressed without any such terms, through a periphrasis or through the implications of an expression. In addition, different normative positions may be called a “right.” Given this, one has to sketch in whatever rough outline the content of what is designated here as a human or fundamental “right”.

A right in the sense explored is a complex bundle of normative positions of a bearer, or many bearers, and the addressee or the addressees of a right. These normative positions or incidents include what is called in standard terminology a right (in a narrower sense), a claim, claim-right, or subjective right of the rights-holder to an action or forbearance and the corresponding duty of the addressee to perform the action or to forebear the action toward the bearer.³⁹ If a person has the right to free speech, the bearer has the claim that the addressee not interfere with her expressing herself and the addressee, e.g. the state, has the duty to forebear interference. This is a necessary connection. There are no claims without duties, though there are morally good acts which are not normative correlatives of claims of the patients of the acts, for example, in the case of an action that is supererogatory. The addressee of a right is permitted but not obliged to use the normative position he enjoys. He enjoys a privilege to do or not do so.⁴⁰ The bearer of a right to free speech can express herself or not, for example. The addressee has no normative claim against the bearer to do one or the other. A right opens a normatively protected space for

³⁸ Cf., e.g., N. Chomsky, *Language and Thought*, 1993, p. 23f. It is worth noting that v. Humboldt underlined the reality of universal understanding, W. v. Humboldt, *Über die Verschiedenheit des menschlichen Sprachbaus*, in: idem., *Werke*, 2002, Bd. III, p. 158f.

³⁹ W. N. Hohfeld, *Fundamental Legal Conceptions as Applied in Judicial Reasoning*, ed. D. Campbell/P. Thomas, 2001, p. 12f.

⁴⁰ Ibid., p. 14ff.

the exercise of the discretion of the bearer,⁴¹ defined by the non-existence of a duty having a content precisely opposite to that of the bearer's privilege.⁴²

Rights in this sense can contain but are not the same as normative powers to change the normative position of others. Such powers are sometimes understood as rights, too. The normative ability embodied in a power is correlated to the liability of the patient of the exercise of this power.⁴³ It is opposed to an immunity of the patient to such a power, which implies the disability of the agent to effect such normative changes.⁴⁴

Bare privileges are imaginable: Their normative force is weak and consists in the possibility to act or not to act in a certain way without violating the rights of another person, while at the same time the hindering of the exercise of the privilege does not violate a normative position of the bearer of the privilege.⁴⁵ Fundamental and human rights are no such bare privileges. The permission to do or not to do something goes along and is buttressed by a claim of the bearer against the addressee and a respective duty to do or not to do something, depending on the nature of the right. A legal right is, in addition, enforced by the institutions and the sanctions of the law. Rights are consequently understood below as such a cluster of four normative positions: the claims and privileges of the bearer, the duties and the no-rights of the addressee.⁴⁶

Such moral or legal rights can have a power as content, e.g., the freedom to contract the power buttressed by privileges and claims of the bearer and duties of the addressee to create contractual duties or, in the case of the right to property, the freedom to relinquish ownership (within the limits of other legal norms). This normative position can be accompanied by immunities, e.g., not to be obligated by contracts unless the agent has agreed.

The bearer or bearers of a right vary according to the right concerned. It can be one single individual who has a specific right, e.g., in contractual relations. A group of people, e.g., the residents or citizens of a country, can enjoy the

⁴¹ This is the truth of the so-called will theory, cf. F. v. Savigny, *System des heutigen römischen Rechts*, Bd. 1, 1840, § 4; B. Windscheid, *Lehrbuch des Pandektenrechts*, 9th ed., 1906, Bd. 1, § 37; H. L. A. Hart, *Natural Rights*: J. Bentham and John Stuart Mill, in: idem., *Essays on Bentham*, 1982, p. 80ff. The assumed opposition to the interest theory, cf. R. v. Jhering, *Geist des römischen Rechts*, Teil III, 1924, p. 337ff, is perhaps the product of an incomplete conception of subjective rights: The protection of autonomy may be one of the interests such rights serve and is not without limits, which was the core of Jhering's concern.

⁴² W. N. Hohfeld, *Fundamental Legal Conceptions as Applied in Judicial Reasoning*, ed. D. Campbell/P. Thomas, 2001, p. 14.

⁴³ Ibid., p. 21ff. ⁴⁴ Ibid., p. 28ff.

⁴⁵ Ibid., p. 16; J. J. Thomson, *The Realm of Rights*, 1990, p. 46ff; R. Alexy, *Theorie der Grundrechte*, 1985, p. 203ff.

⁴⁶ For an expression of these relations with the means of deontic logic cf., e.g., R. Alexy, *Theorie der Grundrechte*, 1985, p. 171ff.

same rights. In the case of human rights properly speaking, the bearers of these rights are all humans because of no other further characteristic than the humanity of the person concerned.⁴⁷ Fundamental rights are understood here as encompassing human rights and other rights of central importance that are, however, not granted to all persons, often due to good reasons. The guarantee of human dignity is in many legal systems a human right and, due to its importance, at the same time a fundamental right. The right to vote is universally restricted to citizens or long-term residents of a state or other state-like structures like the EU, and there are evident reasons for that. Still, the right to vote in a *specific* community is of particular importance for a person and thus a fundamental right, though it is not a human right of all human beings to vote in this specific community. The right to participate at all in *some* community is, however, a human right.⁴⁸

Human and fundamental rights can be moral or legal. The relation of law and morality is notoriously contested, and this is no different for the case of human and fundamental rights. A realistic legal hermeneutic teaches the lesson that legal rights are not wholly independent from moral rights.⁴⁹ One will hardly be able to delineate the scope of any fundamental legal right without taking recourse to a principled account of what this particular fundamental right and fundamental rights in general are about. Legal interpretation and application of fundamental rights are necessarily guided by a theory of fundamental rights.⁵⁰

Lawyers and legal theorists sometimes think that they can purify the law of nonlegal influences, especially ethical ones, not least because ethics are regarded as intrinsically contentious, subjective, and thus detrimental to the political point of the law to establish an authoritative order based – in a democracy – on common grounds.⁵¹ This attempt cannot succeed. Especially open-textured norms like human rights are in need of interpretation and

⁴⁷ There is an intense discussion about animal rights. Nothing in these remarks has any direct bearing on the question of which normative status animals enjoy. This is simply a different question.

⁴⁸ Cf. Art. 21 UDHR.

⁴⁹ Cf. M. Mahlmann, *Elemente einer ethischen Grundrechtstheorie*, 2008; R. Alexy, *Theorie der Grundrechte*, 1985; Dworkin has recently even argued that law should be conceptualized as a sub-branch of (political) morality, cf. R. Dworkin, *Justice for Hedgehogs*, 2011, p. 405.

⁵⁰ Cf. M. Mahlmann, The Dictatorship of the Obscure? Values and the Secular Adjudication of Fundamental Rights, in: A. Sajó/R. Uitz (eds.): *Constitutional Topography: Values and Constitutions*, 2010, p. 343ff.

⁵¹ The classic example is H. Kelsen, *Reine Rechtslehre*, 2nd ed., 1960. On the discussions on exclusive and inclusive positivists, cf. W. J. Waluchow, *Inclusive Legal Positivism*, 1994; J. Coleman, *The Practice of Principle*, 2001; J. Raz, *The Authority of Law*, 1979, p. 49f.

concretization. Interpretation will, by hermeneutical necessity, take recourse – explicitly or implicitly, knowingly or unknowingly – to such more or less reflexive theories of fundamental rights that have – among others – an ethical dimension. Whatever you think about the jurisprudence of the European Court of Human Rights on the absolute prohibition of torture and its interpretation of Art. 3 ECHR in this respect, the argument for or against this interpretation will include, whether one wants it or not, intricate ethical arguments about the absolute or relative value of human life, the existence and scope of human dignity, and the competing importance of other values, for example, the rights of third parties in cases of torture used to prevent harm to other persons as in the leading case of the ECtHR on the matter.⁵² In addition, a catalogue of fundamental rights is not justifiable without ethical considerations, as these are the ultimate sources of normative justification.

A right can be addressed to a single individual, a plurality of addressees, or everybody. Rights can be relative or absolute. Legal human rights are directed against public authority – national, supranational, or international. Depending on the state of development of the respective legal system, they have a direct or indirect horizontal effect as well, thus obligating private persons. Positive obligations have the same normative effects, as widely accepted around the globe, with some exceptions.⁵³ Moral human rights are directed to private actors. Nigerian peasants, for example, have a moral right that their environment not be destroyed by the action of private companies producing oil. Moral rights can be directed against public authorities as well. Though one may not have a legal right to build a minaret in Switzerland as the law stands (though the case is far from clear), there are rather plausible grounds to believe that there is a moral right of believers to determine the shape of sacral buildings, within the framework of the general rules, such as building safety, and that this moral right is illegitimately violated by the existing legal ban on minarets.

The content of human rights catalogues varies, often in crucial details. But there is a cluster of central positions that are the basic elements of fundamental rights: dignity, life, bodily integrity, liberty, equality, and solidarity. These are abstract concepts, contentious in detail but not without *prima facie* content. In addition, developed codes include other rights, derived from the general *telos* of human rights catalogues, to serve and protect these core contents, for example, social rights, political rights, rights to institutions relevant for the application of

⁵² ECtHR (Grand Chamber), appl. no. 22978/05, 1 June 2010, Gäfgen v. Germany.

⁵³ Cf. *DeShaney v. Winnebago County* 489 US 189 (1989) (US).

rights, for example a judicial system, and so on. The Universal Declaration is a prime example of this.⁵⁴

These remarks are cursory and leave out many qualifications of rights, especially those of developed legal systems. But they suffice for the limited purpose at hand.

IV THE HISTORY OF RIGHTS

After these clarifications, one can address the next question: How did the idea of human rights evolve? What are the lessons taught by this historical trajectory for the question pursued? Given the analysis above, human rights are in this context best understood as an intricate web of normative positions of all human beings: They—privileges and the absence of duties of the bearer toward others, claims and related duties obligating everyone at least as moral—a normatively circumscribed and protected space for individuals. They secure with these normative means human dignity, life, integrity, freedom, equality, and subsistence in morality and in law.⁵⁵

Looking at contemporary debates, one can observe that historical analysis is prominently used in controversies about human rights with the critical intention of unveiling the dark history of rights. Human rights are taken not to be universal aspirations of humankind and “door openers to closed societies”⁵⁶, challenging illegitimate authority and empowering the weak, but rather

⁵⁴ Cf., e.g., for social rights Art. 22–27 UDHR; for political rights Art. 21 UDHR; for institutional rights Art. 10, 28 UDHR.

⁵⁵ In view of some controversies in current debates: Human rights are not restricted to legal provisions of international law. On the one hand, they designate a normative position independent of positive law (in natural law theory or in contemporary theory of morality) of the content described. On the other hand, law has developed different tools to protect such rights: First of all, constitutional rights that extend their protection not only to citizens but all humans, irrespective of citizenship or residence rights. In contemporary, well-developed constitutional systems this is, e.g., the case for the central right of human dignity. Another tool is traditional international law. Proto-state-like structures like the EU and its supranational legal order provide for similar protection. For some details on these interwoven fields of law cf. M. Mahlmann, *Human Dignity and Autonomy in Modern Constitutional Orders*, in: M. Rosenfeld/A. Sajó (eds.), *The Oxford Handbook of Comparative Constitutional Law*, 2012, p. 370ff. For an example of a deficient understanding of the technical meaning of human rights cf. S. Moyn, *The Last Utopia*, 2010, identifying them with international norms limiting state sovereignty.

⁵⁶ L. Wingert, *Türöffner zu geschlossenen Gesellschaften*, in: R. Elm (ed.), *Ethik, Politik, Kulturen im Globalisierungsprozess*, 2003, p. 392ff. This function as a “door opener” is not only a political one, but in addition, as Wingert rightly argues, has an epistemic dimension as well: The idea of human rights helps to identify violations of basic normative positions. This is an important claim: Human rights understood in this sense are a heuristic tool for the discovery of injustice.

shrewd plots of partisan politics and political religion.⁵⁷ This is the challenge of the historical, genealogical human rights revisionism mentioned above. The contingent origin of human rights is not just a neutral historical fact, but their dark heritage shows that they are not a “last utopia” but are ultimately delegitimized by their reactionary, religiously biased origin, despite all later transformations, which are ultimately doubtful in their effects.⁵⁸

This perspective may appear surprising because a historical development as such cannot justify or delegitimize a normative institution: The facticity of a historical trajectory provides no normative reasons. The course of history is one thing, the justification of the products of history quite another, not least in the case of human rights.

But given the current debates and the underlying thought that genealogy may indeed be relevant for justification, it is helpful to pause for a moment and ask: Are these voices perhaps right? Are human rights in fact contaminated by their dark origins?

The first important step in any enquiry in human rights assessing the merits of such claims is not to commit the methodological fallacy of looking for words or terms like *rights*, *Recht*, *ius*, and so on but for the idea, the meaning of such historically and currently employed terms. What this idea is about has been outlined above, and it has at least two central dimensions, the *formal concept* of a right and the *particular content* of what is regarded as fundamental or human rights. A proper history of human and fundamental rights must look at both dimensions and see how and when they overlap.

Consequently, if one thinks about the history of ideas of human rights, one has to search for explicit and, the deeper one digs, most probably implicit expressions of this idea. This move is central to avoiding a second fallacy: the fallacy of intellectual and often cultural, and more precisely Western elitism.

⁵⁷ Cf. S. Moyn, Personalism, community, and the origins of human rights, in: S.-L. Hoffmann/S. Moyn (eds.), *Human Rights in the Twentieth Century*, 2011, p. 85ff, 87: “Human rights need to be closely linked, in their beginnings, to an epoch-making reinvention of conservatism”; idem., *The Last Utopia*, 2010, p. 47: “After a few years had passed, the meanings the idea of human rights had accreted were so geographically specific and ideologically partisan – and, most often, linked so inseparably to Christian, Cold War identity – as to make the fact that they could return later in some different guise a deep puzzle,” *ibid.*, p. 54, 74ff.

⁵⁸ Cf. S. Moyn, *The Last Utopia*, 2010, p. 225ff on the intrinsic limits and burdens of the human rights idea. On the case study of human dignity, and a critique on similar grounds, with the conclusion that the concept is useless, idem., The Secret History of Constitutional Dignity, in: C. McCrudden (ed.), *Understanding Human Dignity*, 2013, p. 95ff, 95, 111; idem., The Continuing Perplexities of Human Rights, *Qui Parle* 22 (2013), p. 95ff, 107ff underlines his skepticism, with some qualifications, given that so far there is nothing much better. On somewhat more positive notes, idem., *Human Rights and the Uses of History*, 2014, p. 135ff. On religious bias, cf. idem., *Christian Human Rights*, 2015.

There are even today people who lack any words for human rights, but do not lack the central idea, both as to the normative form and as to their essential content. An illiterate woman being mistreated and perhaps raped by Boko Haram forces may not be able to express the normative claims she thinks she justifiably has in Hohfeldian terms, nor their content with the technical language of Kantian conceptions of dignity, but she may, quite rightly, have the idea, in whatever obscure form, that this is not right, that she should be free to be left alone, to have her body and inner self not harmed by such acts, that she, as a human being, can rightly demand not to be treated like this, neither by state forces nor by private parties and that everybody has to abstain from hurting her in these pernicious ways. It would be a gross and appalling failure of any theory of human rights to not take notice of this fact.

The same is obviously true of historical perspectives. It is, therefore, not good enough to enter terms in search engines, to paste digitalized historical texts for occurrences of terms like *rights*, *Rechte*, *ius* and so on to truly write an intellectual history of the idea of fundamental and human rights.⁵⁹ One has to look for this idea, or at least basic elements of it, and not only in the canonical texts of high culture but in the social practices, and not least the struggles of ordinary human beings, who in the course of history on many occasions have manifested their belief and the reasons for this belief that they enjoy a particular normative position, a justifiably claimable permission to do or not do something and to have central goods of life protected. There are other sources as well, such as the aesthetical self-representation and self-appropriation of human existences by its subjects in art. A history of human rights without at least a sense of what has been expressed, for example, about human dignity in the aesthetical sphere will again miss a crucial aspect of the development of such ideas. Or is, to make it concrete with some examples, Odysseus' encounter with the ephemeral and fading shadow of his mother in the underworld and what it tells us about the meaning of mortality, the suffering of Euripides' women of Troy, the mourning of the seated, medieval terracotta figures of the Niger delta, Scheherazade's songs against death, Basho's adrift farewell note,⁶⁰ or the rebellious defiance of the upright figure with cicatrices in front of the firing squad in Goya's *El 3 de mayo de 1808 en Madrid* entirely irrelevant if one asks whether dignity is a property of the human condition?

⁵⁹ As E. Hunt, *Inventing Human Rights*, 2007, p. 230, note 5, reports to have done. There is nothing wrong with such research; on the contrary, it is very useful if one does not mistake it for a comprehensive search for what human rights history should be interested in.

⁶⁰ "Ailing on my travels / yet my dream wandering / over withered moors," translation G. Bonnas/A. Thwaite, *The Penguin Book of Japanese Verse*, 1998, p. 106.

Even if one restricts one's view to the history of thought as expressed in canonical texts, it is clear that the traces of the idea of human rights are manifold. Three examples from very different epochs may suffice: Antiquity, Early Modernity and the Enlightenment. Important contributions to the historical analysis of rights have not missed the point that one has to look beneath the surface of the use of terms to seriously engage in the issue.⁶¹ As a result, it has been argued – and quite plausibly so – that one can find an implicit statement of rights in central theories of justice in antiquity not because certain words were found in texts but because a proper, substantive, and of course difficult analysis of these theories and what they entail makes such interpretation convincing.⁶² To be sure, rights in such theories were not Universal Declarations in Greek prose, and in many respects quite contrary to a plausible set of human rights⁶³ but nevertheless something relevant and interesting from the point of view of the history of ideas. This is the case not only as far as the formal category and its use are concerned, but for the history of the incremental drawing of their substantial contours as well: for example, if one thinks about Socrates' respect for human autonomy⁶⁴ and the attitude toward others and their worth that it entails.

The natural law tradition in the European context provides many other examples. Take Grotius. He is widely regarded as restating Scholastic reflection on natural law with its roots in patristic thought and the Stoa⁶⁵ and transforming it into the conception of the modern natural law tradition.⁶⁶ He had an

⁶¹ Cf. G. Vlastos, Rights of Persons in Plato, in: idem., *Studies in Greek Philosophy*, Vol. II, *Socrates, Plato, and Their Tradition*, 1995, p. 124 on the irrelevance of the linguistic fact for a theory of rights that to express the idea of a right various ancient Greek terms were employed.

⁶² G. Vlastos, Rights of Persons in Plato, in: idem., *Studies in Greek Philosophy*, Vol. II, *Socrates, Plato, and Their Tradition*, 1995, p. 104ff; and the semantic analysis idem., Social Justice and the Polis, in: idem., *Studies in Greek Philosophy*, Vol. II, *Socrates, Plato, and Their Tradition*, 1995, p. 70ff.

⁶³ As G. Vlastos correctly observes, the core problem of Plato's theory of justice is that the citizens of the polis are ultimately not regarded as an end-in-themselves, idem., Social Justice and the Polis, in: idem., *Studies in Greek Philosophy*, Vol. II, *Socrates, Plato, and Their Tradition*, 1995, p. 91.

⁶⁴ For example, implying moral autonomy of his partners in dialogue, cf. G. Vlastos, *Socrates*, 1991, p. 44.

⁶⁵ On the scholastic and antique roots cf., e.g., T. Irvin, *The Development of Ethics*, Vol. II. 2008, p. 99.

⁶⁶ Cf., e.g., C. Thomasius, *Fundamenta juris naturae et gentium*. Editio Quarta. 1718, p. 3, § 1 on natural law theory: "Uti enim Grotius hanc utilissimam disciplinam pulvere scholastico commaculatam & corruptam, ac tantum non exanimaram primus iterum suscitavit ac purgare incepit"; or K. Haakonsson, Hugo Grotius and the History of Political Thought, in: *Political Theory* 13 (1985), p. 239ff, 239.

analytically sophisticated idea of rights of human beings⁶⁷ fundamentally not restricted to certain groups of beings,⁶⁸ though his theory did not lack elements that could buttress the imperial ideology of European states that inflicted much harm and suffering on the people subjected to the rule of European powers.⁶⁹ This example is of some importance: Grotius' work formed for centuries the textbook not only of the legal but of the ethical thought of the intellectual quarters of the European cultural context.⁷⁰ That finally the idea of fundamental and human rights was prominent in the Enlightenment and its immediate precursors needs no substantial argument.

In recent times, an increasing effort has been made to examine critically the biases and limitations of classical elements of legal thought and to do justice to other traditions of practical philosophy outside of the so-called Western tradition.⁷¹ There is good reason to pursue this effort, which adds yet another dimension of complexity to the scope of the history of human rights.

Any historical account of human rights has to pay close attention to these matters – to the richness of thought on rights in history, the need for analysis of theories for implied ideas instead of the search for words, the importance of a keen awareness of the many forms of human expression and their often highly significant content beyond theoretical thought, the necessary respect for the profound ethical implication of human practices beyond elite circles, not least of human struggles, and the necessary abandonment of Eurocentric or otherwise parochial perspectives. In light of all this and what it tells us about the history of human rights, the thesis that the idea of human rights and its significance is the recent product of activities by Amnesty International (admirable as they are),⁷² by Jimmy Carter⁷³ or by Catholic Personalism⁷⁴

⁶⁷ Cf. H. Grotius' notion of rights, H. Grotius, *De Iure Belli ac Pacis Libri Tres*, ed. J. Brown Scott, 1995, I, IV, V, XVII.

⁶⁸ Ibid., I, IVff.

⁶⁹ Cf. his opinions in idem., *De jure praedae commentarius*, ed. H.G. Hamaker, 1868. On the context of Dutch colonialism R. Tuck, *The Rights of War and Peace*, 1999, p. 79ff.

⁷⁰ R. W. Lee, Hugo Grotius, in: *Proceedings of the British Academy* 16 (1930), p. 219ff, 267: *De jure belli ac pacis* "supplied the nations, particularly the protestant nations, of Europe with what they wanted – a rational theory of international relations emancipated from theology and the authority of churches. It was well adapted to be the textbook of the New Europe (a congeries of independent powers) to which the Peace of Westphalia had set its seal."

⁷¹ E.g., A. Sen, *The Idea of Justice*, 2009, p. 37ff.

⁷² Cf. S. Moyn, *The Last Utopia*, 2010, p. 3ff: "The drama of human rights, then, is that they emerged in the 1970s seemingly from nowhere."

⁷³ Cf. ibid., p. 217: "During the Carter administration, to which it clearly owed its newfound public role, the human rights movement generally treated government as an ally."

⁷⁴ S. Moyn, Personalism, community, and the origins of human rights, in: S.-L. Hoffmann/S. Moyn (eds.), *Human Rights in the Twentieth Century*, 2011, p. 85ff.

appears to be strikingly misleading. The history of human rights is one of “continuity *and* discontinuity”,⁷⁵ and the idea of human rights as outlined above has certainly not been explicitly formulated since the beginning of time in all cultures. This is not least because human history is a long journey toward human self-understanding, a winding, tragic course with early advances and epochs of regress, slowly fathoming humans’ most important concerns and vulnerabilities, finally forming a concept of humanity that is not biased and fragmented by ideologies of exclusion based on sex, skin color, cultural origin, and the like. There are, however, many traces of the idea discoverable in history, in the higher echelons of famous thought and art of many cultures, but also in the backyards of history, where the forgotten, often downtrodden and quickly forgotten dwell and are often repressed and still show, in acts of rebellion and daily uprightness, what rights, though denied, are truly theirs.

Thus, the roots of human rights in history are deep. They are not the monopoly of just one culture, system of thought, religion, or political agenda. To be sure, human rights have been and currently are abused for political purposes. This seems to be the fate of any great human idea. But the idea is not necessarily delegitimized by the abuse of its political foes. Human rights as embodied in law are equally not just the innocent offspring of beautiful moral minds. They are the product of many political forces including those that were sometimes far from any serious attachment to the idea of human rights.⁷⁶ Again, this comes as no surprise given the way human institutions are formed in history. But human rights are not necessarily contaminated by these origins if they are otherwise justified. Overlooking this, one fails to provide a crucial element of human rights history: “a historically convincing account of the normative power of the human rights idea.”⁷⁷

The interesting lessons of the history of ideas thus raise a question that is very different from the one discussed by current historical human rights

⁷⁵ Cf. C. McCrudden, Human Rights Histories, *Oxford Journal of Legal Studies* 35 (2015), p. 179ff, 181. S. Moyn, The Continuing Perplexities of Human Rights, *Qui Parle* 22 (2013), p. 95ff, 96ff, admits this. He defends himself with the remark that he is interested in conceptual history, not the history of an idea, *ibid.*, p. 98. This is on all accounts not what his historical account is about: It does not just trace the use of terms (if that is meant by conceptual history as opposed to a history of ideas), but what he takes as the emergence of an apolitical, moralist, and impoverished utopia. On the logic of inclusion of formerly excluded groups L. Hunt, *Inventing Human Rights*, 2008, p. 146ff.

⁷⁶ Cf. the famous example of Jan Smuts and the inclusion of the term “dignity” in the UN Charter, M. Mazower, *No Enchanted Palace*, 2009, p. 28ff.

⁷⁷ C. McCrudden, Human Rights Histories, *Oxford Journal of Legal Studies* 35 (2015), p. 179ff, 203.

revisionism. The question is this: Why did the idea of human rights in one way or another spring up in quite different historical and cultural contexts? Why does it prosper today, not least at grass-roots level, despite its many political foes? What are the deeper reasons for this fact? One answer to this fundamental question is the assumption of a practical understanding in principle shared by all human beings the exercise of which leads them to some insight about rights if they reflect on their existential situation. This is in no way a surprising or strange hypothesis because it is the fundamental assumption of many streams of thought in human history, not only of the evident cases like natural law theory and the theory of moral understanding of the Enlightenment. It is the hidden working hypothesis of the contemporary human rights culture, of millions of people, despite the importance of various relativisms for some academic and political debates. The foundational assumption of the project of human rights is that human rights make sense for everyone. The whole project starts from the idea that any human being endowed with “reason and conscience” is capable of understanding that there are human rights for herself and others – irrespective of culture, upbringing, gender, skin color, and the like. The human rights project is a piece of optimism about the possibility of common human insight: Despite the many obstacles – not least powerful ideologies, incited hatred, the cultivation of moral parochialism in some quarters – it is regarded as possible that humans can come, via perhaps long and painful processes of reflection and cultural change, to understand that such rights exist and that they are worth the care, passion, and sacrifice needed to make them a living thing.

Thus the true answer to the challenge of genealogical, historical human rights revisionism is not the proof of the historical universal presence of human rights in human history. Such proof is neither possible nor necessary. The true answer to the historical, genealogical and normative relativizing of human rights is the reassertion of the strength of the reasons, first, for the validity of human rights and, second, for the assumption of the reality of a fundamental and universal faculty of human beings for moral cognition that provides the epistemic access to the idea of human rights for everyone. This human faculty of moral insight is not just the privilege of some philosopher kings, of some special peoples, cultures, or religions. It is not the privilege of whites and men, and it is not the privilege of just one time. The history of the idea and practice of human rights is from this point of view the history of constant renewed approaches to this great idea, in very different forms, historical, social, cultural, and religious contexts, often implicit, always fragmentary, tentative, and imperfect. Our time is just one chapter of this history. And, to be sure, other times will discern better the limits of our understanding

and practice of fundamental rights. There is no reason to assume that we will fare any better than other times, that we grasp fully what is so difficult to develop and even harder to fill with real life.

The critique of a widely discussed challenge to human rights consequently puts two questions on the table, which are the real core of the matter: First, are human rights justifiable, and if so, how? Second, what are the conditions for the epistemological and ontological plausibility of such a theory of justification? What kind of theory of normative cognition does it presuppose?

Historical or even historicist accounts of human rights thus offer no escape route from the theory of the justification of human rights and their epistemological and ontological aspects. This shows that the review of the merits of historical human rights revisionism is no unjustified digression but, on the contrary, leads right to the central topic of these remarks, the question of the relationship between human thought and the idea of human rights. More precisely, it raises the question whether the idea of a universal faculty of moral insight, of practical reason, if you will, shared by all human beings, and thus an epistemological alternative to various forms of perspectivism,⁷⁸ can make any sense for contemporary thought.⁷⁹

The precondition for answering this is a theory of justification of human rights beyond the particular political or religious strategies that hoped to use the idea of human rights for their purposes. Such a theory tells us which reasons justify human rights. Only if that is clarified can the question be asked whether a certain structure of the human mind has any importance for the project of justification, not least to delegitimize it, as is claimed by neuroscientific, neo-emotivist human rights skepticism to be discussed afterward.

V THE JUSTIFICATION OF RIGHTS

1 Variety of Justifications

Are human rights justified? To answer this question, one has to distinguish between explanatory and justificatory theories, though sometimes explanations and statements about the legitimacy or illegitimacy of human rights are

⁷⁸ On perspectivism, cf. F. Nietzsche, *Jenseits von Gut und Böse*, in: idem., *Kritische Studienausgabe*, G. Colli/M. Montinari (eds.), Bd. 5, S. 12. Historical revisionism is not necessarily wedded to perspectivism: It could claim that human beings do share a fundamental perspective on the world through time that did not, however, produce a concept of human rights till the 1970s.

⁷⁹ That this is not obvious is illustrated by Habermas' remark that subjective reason is "zerbrochen", in pieces, idem., *Faktizität und Geltung*, 1992, p. 17.

combined, for instance, in the genealogical revisionism discussed above, but in other approaches as well that aim to defend the idea of human rights.⁸⁰ Explanatory theories tell us about the causes of the existence of human rights. Many causes are imaginable – social, historic, political, anthropological, cognitive. None of them as such has any bearing on the justification of human rights, as the latter are ultimately not based on any causal account of the origin of human rights, but on the normative grounds of their legitimacy – a topic to which we will return.

There are various justificatory theories of human rights, some with deontological, some with consequentialist leanings, some from the quarters of virtue ethics. *Social functionalist theories* argue, for instance, that human rights are not only explained but justified by their function for successful social integration by keeping “the future open” for the reproduction of an autopoietic social system.⁸¹ Arguments for human rights from the point of view of the *economic analysis of law* point out that certain forms of human rights increase efficiency.⁸² Such analysis may be informed by *behavioral economics* and research on psychological heuristics, framing patterns and biases like risk aversion within prospect theory,⁸³ factoring a deontological threshold into the cost-benefit analysis.⁸⁴ *Contractualism* justifies human rights by the assumed consent of the parties of the contract,⁸⁵ *discourse theory* by the consensus reached after a deliberative process under the conditions of the ideal speech situation, where not might but the “zwanglose Zwang des besseren Arguments,” the “unforced force of better arguments,” reigns.⁸⁶ This deliberative process

⁸⁰ H. Joas, *Die Sakralität der Person*, 2011.

⁸¹ Cf., e.g., N. Luhmann, *Die Gesellschaft der Gesellschaft*, 1997, p. 1094ff.

⁸² Cf., for an example of such reasoning from the analysis of constitutional rights, R. Posner, *Economic Analysis of Law*, 9th ed., 2014, p. 978: “A search (or seizure) is reasonable if the cost of the search in privacy impaired (B) is less than the probability (P) that without the search the target of the search cannot be convicted or otherwise rendered harmless . . . , multiplied by the social loss (L) if he eludes punishment.” For an analysis of torture, *ibid.*, p. 984, arguing that under normal circumstances, torture is regularly too costly, but is efficient in the case of prevention of terrorist attacks: “The cost-benefit analysis of coercive interrogation would be dramatically altered if for example the interrogation concerned a terrorist plot and the person interrogated – a peripheral figure in the plot but a possessor of vital information – faced no criminal punishment but merely deportation as an illegal alien, continued surveillance, or a warning.”

⁸³ An interesting example is the reconstruction of the debate about civil and political rights on the one side and social and economic rights on the other from the perspective of them being framed as losses or gains, cf. E. Zamir, *Law, Psychology and Morality*, 2015, p. 143ff.

⁸⁴ On threshold deontology E. Zamir/B. Medina, *Law, Economics and Morality*, 2010, p. 41ff.

⁸⁵ E.g., J. Rawls, *A Theory of Justice*, 1971; T. M. Scanlon, *What We Owe to Each Other*, 1998.

⁸⁶ J. Habermas, *Faktizität und Geltung*, 1992, p. 138.

realizes the foundational right to justification⁸⁷ and enables the reflexive self-endorsement of everybody's subjectivity without domination.⁸⁸ *Interest or need theories* hold that some interests or needs of human beings are of such importance that they give rise to human rights.⁸⁹ Others argue that human rights secure the fundamental conditions for pursuing a good life.⁹⁰ *Teleological personhood theories* based on an "expansive naturalism" justify human rights because they are a precondition for human agency.⁹¹ *Interpretationism* holds that the ascription of rights is the result of the best uncircumventably normative interpretation of concepts that are not criterial or of natural kind but are interpretative, including pivotally human dignity as self-respect, authenticity and respect for the humanity of others.⁹²

This cursory list is not conclusive and scratches only the surface of the sophisticated thoughts underlying these approaches.⁹³ Considering these examples, at least in some important respects, however, helps to clarify what elements a theory of human rights has to contain. Social functionalist theories lack, for example, a crucial normative element, not least because some of them misunderstand the function of human rights. Human rights are about protecting central individual goods irrespective of whether this is functional or not for society. Free speech can be used for socially quite dysfunctional purposes. In any case, human rights are instruments to protect persons, not instruments to protect the functioning of society as such. Rights formulate, in addition, yardsticks for the aims of society – whatever they are, such aims must not violate human rights in the first place. As a consequence of this misconception, functionalist theories do not specify the normative principles

⁸⁷ Cf. R. Forst, *Das Recht auf Rechtfertigung*, 2007.

⁸⁸ Cf. K. Günther, *Anerkennung, Verantwortung, Gerechtigkeit*, in: R. Forst/M. Hartmann/R. Jaeggi/M. Saar (eds.), *Sozialphilosophie und Kritik*, 2009, p. 269ff, 269, 286f.

⁸⁹ On interest theories, cf. J. Raz, *The Morality of Freedom*, 1986, p. 166: "Definition: 'X has a right' if and only if X can have rights, and, other things being equal, an aspect of X's well-being (his interest) is a sufficient reason for holding some other person(s) to be under a duty. Capacity for possessing rights: An individual is capable of having rights if and only if either his well-being is of ultimate value or he is an 'artificial person' (e.g., a corporation)." On need theories, cf., e.g., D. Miller, *Grounding Human Rights*, *Critical Review of International Social and Political Philosophy* 15 (2012), p. 407ff, 422. On a critique of need theories and in defense of interest theories, see J. Tasioulas, *On the Foundations of Human Rights*, in: R. Cruft/S. M. Liao/M. Renzo (eds.), *Philosophical Foundations of Human Rights*, 2015, p. 63ff. Tasioulas formulates four very plausible justificatory stages of interest theories, *ibid.*, p. 50f, including possibilities and burdens into a threshold criterion.

⁹⁰ Cf. S. M. Liao, *Human Rights as Fundamental Conditions for a Good Life*, in: R. Cruft/S. M. Liao/M. Renzo (eds.), *Philosophical Foundations of Human Rights*, 2015, p. 79ff.

⁹¹ J. Griffin, *On Human Rights*, 2008, p. 32ff. ⁹² R. Dworkin, *Justice for Hedgehogs*, 2011.

⁹³ Cf., e.g., logically generating rights from human action, A. Gewirth, *The Community of Rights*, 1996, p. 13ff, or the "practical conception" of C. Beitz, *The Idea of Human Rights*, 2009, p. 102.

underlying the allocation of rights and for determining the social aims these rights serve, including the normative calibration of what societies are there for.

The efficiency-based rights-accounts have comparable shortcomings: They miss the central points of human rights to provide normative principles beyond efficiency. Rights formulate limiting conditions for any efficiency regime and therefore presuppose a justification that transcends the limits of a cost-benefit analysis.⁹⁴ Behavioral law and economics is in need of such principles as well: Skewed decisions due to heuristics, framing effects, or biases are important elements of a realistic theory of human decision-making. The picture of bounded rationality, however, leads to the problem of what principles should be guiding human decision-making instead, not least which normative principles are to be decisive to overcome the shortcomings of common human decision-making.⁹⁵ Contractualism leaves the questions of the normative foundations of the contract open, more precisely: Why is it legitimate that human beings are only bound by obligations that can be imagined as accepted by free and equal persons?⁹⁶ Where do the noncontractarian preconditions of contractualism stem from? Discourse theory correctly highlights the irreplaceable individual autonomy and subjectivity that are non-negotiable normative yardsticks of legitimate normative content.⁹⁷ Notwithstanding, it faces a related problem: A discourse in the technical sense is a normatively loaded enterprise. One may discuss the reach of the minimal ethics embodied in communicative action. Whatever its scope may be, it does not include the demand for inclusion of anybody in discourse patterns in the first

⁹⁴ To take the example mentioned before: The point of limiting governmental searches and seizures is the protection of human freedom and autonomy. This idea can justify much stricter limits than conceived in Posner's formula, see above, Fn. 82. This is even more evident for the crucial example of torture, where the dignity of a person justifies an absolute prohibition of such practices. A cost-benefit analysis of the kind imagined by Posner, *ibid.*, opens the door wide to practices that are not reconcilable with human rights.

⁹⁵ The debate about libertarian paternalism and its limits is exactly about this question, cf. R. Thaler/C. Sunstein, *Nudge*, 2008. For a critique J. Waldron, *It's All for Your Own Good*, *The New York Review of Books* 61/ 15 (2014); C. McCrudden/J. King, *The Dark Side of Nudging: The Ethics, Political Economy, and Law of Libertarian Paternalism*, in: A. Kemmerer/C. Möller/M. Steinbeis/G. Wagner (eds.), *Choice Architecture in Democracies: Exploring the Legitimacy of Nudging*, 2016, p. 75ff. Theories that defend and incorporate deontology need a justification of such principles as well. To refer to the conformity with common moral intuitions is important but not sufficient, if the normative reasons for referring to these intuitions are not clarified, cf. E. Zamir/B. Medina, *Law, Economics and Morality*, p. 65.

⁹⁶ This may be true but is not evidently so. Hegel was, for example, horrified by the idea that a contract could be imagined as founding something sublime like a state, cf. G. W. F. Hegel, *Grundlinien der Philosophie des Rechts*, in: *idem.*, *Werke*, E. Moldenhauer/K. M. Michel (eds.), Bd. 7, 1986, § 273.

⁹⁷ That the individual is not substitutable does not mean that the results of moral thought are private; cf. L. Wingert, *Gemeinsinn und Moral*, 1993, p. 290ff.

place and more generally the kind of respect for human beings that human rights embody and that is not dependent on patterns of communication. Respect for other humans and their rights is the ultimate normative precondition for discursive and deliberative practices and civilized *Lebenswelten* (life worlds); discursive practices and civilized *Lebenswelten* can thus not be the ultimate normative foundations of human rights.

Interest and need theories make an important constructive point by underlining that there is no theory of human rights without reference to the goods these rights protect. The same is true for an approach deriving rights from fundamental conditions for pursuing a good life. These theories face, however, a common problem: How do certain interests, needs or fundamental conditions for pursuing a good life give rise to normative claims and privileges on the side of the rights-holder and obligations on the side of the addressee of rights? This is not obvious. All kinds of interests or needs have no normative consequences at all. Why is it different in the case of some interests or needs? There is a gap in the argument between the descriptive proposition that humans have certain (important, existential) interests or needs and the normative proposition that they legitimately have rights to have these interests and needs secured. This is so even if one qualifies the theory by a threshold criterion according to which a precondition for a human right is that the interest protected can possibly be satisfied and the consequences of the right are not too burdensome for others.⁹⁸ If one relies instead of interests and needs on fundamental conditions for pursuing a good life a comparable question arises. The best answer to bridge this gap seems to be this: Humans enjoy fundamental rights because of normative principles that prescribe under which conditions certain goods are normatively relevant and create claims, privileges, and obligations. Principles of justice are of central importance here, as will be outlined in more detail below.

Agency theories presuppose that agency is of such importance that it should be protected by rights. That is certainly correct. But where does the right to agency stem from? Why is the importance of something an existence condition for a very specific normative position not only of the bearer but of the addressee(s) as well? And are human rights not, in addition, protecting much

⁹⁸ As J. Tasioulas, *Human Dignity and the Foundations of Human Rights*, in: C. McCrudden (ed.), *Understanding Human Dignity*, 2013, p. 291ff, 296, correctly observes: “Nevertheless, reducing human rights to universal interests is a category error. Interests belong to the domain of prudence or well-being, which concerns what makes a life better for the person living it, whereas human rights are moral standards that impose duties on others, where the violation of the duty entails *wronging* someone in particular – the right holder. Our interests, by contrast, can be impaired in all sorts of ways without any *moral* wrongdoing being in the offing, let alone a directed wrongdoing of this specific kind” (emphasis in the original). To solve the problem he introduces the thresholds of possibility and burden, *ibid.*, p. 297ff.

more than human agency?⁹⁹ Finally interpretationism: The leading principle is dignity understood as self-respect and respect for others. This is a very solid principle – but what are the reasons for its defense? What is the answer to the many skeptical voices about dignity?¹⁰⁰

One can learn much from these theories, and no one working in this field should feel very confident that he or she will produce anything as remotely thoughtful as them. Notwithstanding, none of them seems to answer the question of the justification of rights entirely satisfactorily. Therefore, it may be useful to take one step back and to consider anew how one may be able to make some progress in this respect.

2 *Justification and Its Content*

Rights are a normative instrument to protect certain goods of human beings. The term “goods” in the sense employed here is meant to encompass anything that is of value for human beings.¹⁰¹ The details of the scope of human rights are highly contentious and shaped and refined daily in the sphere of law by the work of thousands of lawyers, court decisions, legislative acts, and deliberation about new normative challenges. Core elements are, however, respect for the particular, supreme, and unalienable worth of human beings, their physical and psychic integrity, freedom, equal treatment, and – more contentiously perhaps – the means for their physical subsistence and, more ambitiously, the minimum material preconditions of a dignified life. Human rights imply the idea of a justified distribution and allocation of central goods like respect and freedom. Consequently, any theory of human rights will have to contain at least two things: It has to take a stand – explicitly or implicitly – first on a theory of human goods and second on normative principles that give rise to the particular rights that persons possess and that justify the distribution of goods realized by rights.

Three theses map the outline of what may be a promising way ahead to formulate such a justificatory theory of human rights:

First, there is no theory of human goods without substantive anthropology. Human rights protect goods not of any being, but of human beings. There is no justification of the importance of liberty and self-determination without the assumption that human beings are made of such stuff that liberty and self-

⁹⁹ This leads to the problem of Griffin’s rather restrictive account of the bearers of rights, *idem.*, *On Human Rights*, 2008, p. 83ff, excluding infants and the seriously disabled, assuming that children’s rights are acquired in “stages,” *ibid.*, p. 95.

¹⁰⁰ Cf. on the debate C. McCrudden (ed.), *Understanding Human Dignity*, 2013.

¹⁰¹ A thorough discussion of interests and needs cannot be pursued here. For some comments, cf. J. Tasioulas, *On the Foundations of Human Rights*, in: R. Cruft/S. M. Liao/M. Renzo (eds.), *Philosophical Foundations of Human Rights*, 2015, p. 63ff.

determination are of great worth for them. For ants, the problem of self-determination does not arise, admirable organisms that they are. For human beings it does. Bees do not have a problem with monarchies, human beings do. There is much discussion about the commensurability and interpersonal comparison of the goods of different agents. An important element of the answer to this question is the shared human nature of human beings. To be sure, human beings have widely different preferences, interests, and needs. But some basic, non-trivial things are common, and human rights are an expression of a fundamental understanding of these shared human interests and needs. The project of human rights expresses nothing less than a somewhat noble thesis about the elements of human nature, including the assumption that life, liberty, equality and respect are of such universal importance for human beings that they must be protected by rights. Government, it has been famously argued, is the “greatest of all reflections on human nature”¹⁰² and human rights, by now a central part of any account of civilized government, are part of this reflective enterprise.

Second, any theory of human rights embodies a theory of the historical, social, and political conditions of the realization of human goods. Freedom of speech presupposes the idea that the free exchange of ideas and views does in fact enhance the flourishing of individuals and a community. Dictatorial authoritarianism, on the contrary, presupposes that only narrowly circumscribed expression benefits a community and ultimately the individuals (if its doctrine is not anyway just some cynical ideological device for the subjugation of people). Freedom of faith presupposes that it is not in the true interest of persons that all other religions but the True One have to be suppressed. Again, that is not a trivial assumption. Centuries of violent persecution of believers of other faiths and their intellectual defense testify to that. The deeper one engages with the justification of rights, the more important this aspect of a theory of human rights becomes, because the more argument is needed to show that a particular framing or interpretation of human rights does in fact serve the ends human rights are there to foster.

Third, any theory of human rights has to identify the core normative principles that enter constitutively in the justification of human rights. Plausible candidates are, first, principles of egalitarian justice and, second, principles of human care and solidarity. Principles of justice are the key to the problem why human rights are conceptualized as equal rights. They also provide the answer as to why certain goods – be they interests or needs – have normative relevance. The reason is the connection of justice and rights: Interests or needs are, as such, normatively neutral. The justness of the distribution of goods, both as regards humanity being the criterion of distribution and as regards the interpersonal standard of

¹⁰² J. Madison, in: J. Madison/A. Hamilton/J. Jay, *The Federalist Papers*, 1987, LI, p. 319.

distribution among humans gives rise to the existence of rights, for instance, to equal status and freedom. Justice is a spring of rights.

The second principle is care or solidarity. There are obligations to care for others; this is a core of morality and ethics. These obligations are not boundless, but there is, for example, a duty not to let somebody die who can be helped without compromising overwhelmingly important interests of the agent. Rights protect human goods that are often of existential importance; they can be a matter of life and death. It is therefore a command of human care and solidarity to contribute to the protection of these goods, which gives rise to the rights of others. A consequence is the obligation to strengthen mechanisms of rights protection – by sustaining a legal order, for instance, that protects such rights by one's taxes, by lending political support to international institutions for the protection of human rights or by participating in an NGO committed to improving the political culture of human rights. Alongside justice, duty to care is equally a source of rights.

Real life human rights protection poses many very intricate problems, including questions of what morally legitimate burdens can be imposed on others, *prima facie* and all-things-considered, and of the possibilities and limits of securing particular human goods by rights. Whether or not it is reconcilable with freedom of religion to ban *burqas* in public can only be answered after substantial reflection about the *telos* of this right, the kind of freedom protected, some serious thought about what a society can demand of a person, whether “*vivre ensemble*,” “living together,” is, in fact, a sufficient reason to force a person not to wear a *burqa* (as the ECtHR has argued), the impact of such a ban on the women concerned, whether it liberates them from oppression or drives them deeper into the dungeon of isolation, and so on.¹⁰³ The principles considered do not answer all these questions. They are thus only a part of the full unfolding of human rights, not least in the hard doctrinal work of the law, but a necessary part that will therefore concern us further. Given the topic of these remarks, the central question is: What is the origin of the normative principles identified? What is their epistemological and ontological status? These are the next problems to be considered.

VI RIGHTS AND MORAL COGNITION

1 *Epistemology and Ontology*

The cognition of the justification of a right is achieved by a series of mental acts, by a set of cognitive judgments. Such a judgment has the propositional content

¹⁰³ Cf. ECtHR, app. no. 43835/11, 1 July 2014, *SAS v France*.

“right *x* is justified”, e.g., through expressions like “freedom of speech is worth defending” and so on. This proposition is, in fact, a complex proposition. It implies, as explained, assertions about the usefulness of a certain right like free speech in human communities (for the expression of human personality, the pursuit of truth, the protection of democracy and the like) and, though not always made explicit, anthropological claims about the importance of freedom for human beings. The complexity is increased by the fact that a human right does sometimes appear in very abstract forms.¹⁰⁴ More often than not, however, the norm is stated in more differentiated terms, including a regime of justifications of interferences.¹⁰⁵ That is of crucial importance because this regime of justified interferences determines the true content of the right. The final judgment “right *x* is justified” thus encompasses, in this case, intricate arguments not only about the *prima facie* scope of a right, but also about other values that justify restrictions of this right and the degree to which this is possible. In this context, principles of proportionality have become a core element of modern human rights catalogues, sometimes buttressed by the protection of the essence of a right.¹⁰⁶

The complexity of this judgment will concern us further when we consider some claims of today’s neuroscience on this matter, because overlooking the complexity of this judgment is one precondition, as we will see, for a certain radical neuroscientific critique of human rights.

One element of this bundle of predications is of special concern for normative theory: the predication of the *moral* justification of a right. This is a necessary element of any justification of human rights. There are many considerations that can speak for a human right, but there is no legitimacy of human rights if these rights are not justified from a moral point of view. Nobody would have taken the UN General Assembly seriously if it had stated that to adopt the Universal Declaration made good sense even though it is quite unjust as a catalogue of rights. A human rights catalogue that does not claim to lay down a just and morally appropriate set of norms is not a proper human rights catalogue.

¹⁰⁴ Cf., e.g., as in Art. 1 UDHR or Art. 1 German Basic Law or Art. 7 Swiss Federal Constitution.

¹⁰⁵ There can be horizontal clauses on justified interferences as in Art. 29 II UDHR or in Art. 52 European Charter of Fundamental Rights (European Union, Charter of Fundamental Rights of the European Union, 2012/C 326/02, October 26, 2012). Another technique is a specific limitation clause in the specific right concerned, as in the European Convention on Human Rights (Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, 4 November 1950, ETS 5).

¹⁰⁶ Cf. A paradigm case is Art. 36 Constitution of South Africa. On the essence of a right, cf. Art. 19 II German Basic Law, Art. 36 IV Swiss Federal Constitution.

The predication of the normative justification of a right in this narrow sense by a moral judgment (as an element of a bundle of highly complex justificatory arguments about human rights) must be a principled mental act. Caprice is not justification. The application of normative principles to the case evaluated – the human right under consideration – yields the mental judgment of the moral justification of this right. But what are these principles? And how are these principles used for the justification of norms like human rights justified themselves? These are two of the great questions of practical philosophy that lead to the core of the normative component of the theory of human rights. Accordingly, many principles have been formulated in the history of ideas to pin down at least some core elements of morality, from Socrates' thought that it is better to suffer injustice than to do injustice to Kant's categorical imperative and beyond. Given what has been said about the justification of rights, it is a plausible thesis that principles of equality, and thus of justice, and duties of human care play a crucial role in this respect.

The problems of the content and justification of normative principles guiding moral judgment lead to the next question of interest here. Is the content of these justificatory principles perhaps dependent on the structure of the human mind? Are these principles possibly even in one way or another *determined* by the properties of the human mind? Are the nature and structure of human thought defining factors for the results of reflection in the moral domain – moral judgments? This is, as indicated above, the Cartesian, Lockean, Humean, or Kantian question about the preconditions of the possibility of human moral cognition. Are there such *Verstandesbegriffe*, such “concepts of understanding”, in the moral domain, to use a Kantian term for the building blocks of human theoretical cognition? If so – what is the effect? Are the particular nature and structure of human thought a key to objective, foundational moral insight? Or are they indispensable elements of moral cognition, which leave, however, “morality in itself” (to modify another Kantian idea), the true, objective morality, undisclosed? Is there thus a parallel to theoretical cognition, as human thought, in Kant's view, never grasps the nature of the *Ding an sich*, the thing in itself?¹⁰⁷ Or – a third possibility – do such structures of the mind exist, and do they decisively influence human moral thought, but create nothing but cognitive illusions, the moral equivalent to the Müller-Lyer visual illusion mentioned above?

These questions lead us to the second central challenge to human rights to be considered because a particular stream of cognitive psychology and neuroscience has formulated exactly the latter thesis: It holds that the nature and

¹⁰⁷ I. Kant, *Kritik der reinen Vernunft*, Akademie Ausgabe, Bd. III, p. 16f, 202ff, 207.

structure of the human mind is actually decisive for human moral reasoning but that moral judgment thus determined yields not insight but error, that the idea of human rights is part of this erroneous reasoning and that it is thus not qualified to decisively influence human affairs.

2 *The Neuroscientific Attack on Human Rights: Human Rights and the Mental Gizmo Thesis*

The thesis to be considered accepts as a general framework the dual-process model of the mind that holds that there are two kinds of mental processes, thinking fast and thinking slow, to use the terms with which it has been popularized:¹⁰⁸ Thinking fast means using heuristics, framing operations, or biases for solving everyday problems. These mechanisms are hard-wired in the human brain. Humans cannot but use them and do so intuitively and unconsciously.¹⁰⁹ Thinking slow means reflective rationality that abides by standards of logic.¹¹⁰ Thinking fast works well for many aspects of everyday life but is skewed in important aspects. Human decision-making is thus not fully rational. One can become aware of the factors that skew human rationality like heuristics, framing effects, and biases and overcome their influence by slow thinking, but this is not necessarily the case, because the control system of slow thinking may not do its task.¹¹¹ This model is highly influential not only in psychology but has inspired research in many other fields, not least behavioral law and economics. Its pioneers have not said anything substantial about the place of moral reasoning within this mode. The mental gizmo thesis aims to fill this gap.

The mental gizmo thesis is as follows: Moral cognition is, it says, part of the dual-process mind.¹¹² Deontological judgments are part of fast thinking. There is a mental “gizmo” that yields such judgments, involuntarily, unconsciously, as a fact of the fast, automatic, and emotional operations of the human mind. They are like heuristics or biases: They are useful in certain respects but should be disregarded as general directions of moral judgments because they systematically skew human moral rationality.¹¹³ A prime example of products of the operation of the mental gizmo that cause “moral illusions”¹¹⁴ (in the same way that the Müller-Lyer illusions causes visual illusions) is Kant’s principle of humanity that one should never use other human beings merely as means and not as ends.¹¹⁵ This alone is already important for the topic of human rights,

¹⁰⁸ D. Kahneman, *Thinking, Fast and Slow*, 2011. ¹⁰⁹ Ibid., p. 19ff. ¹¹⁰ Ibid.

¹¹¹ Ibid., p. 39ff. This idea is the birthplace of nudging, the idea that one can systematically exploit these factors for the benefit of others in the framework of a “libertarian paternalism.”

¹¹² J. Greene, *Moral Tribes*, 2013, p. 15, 105ff. ¹¹³ Ibid., p. 132ff. ¹¹⁴ Ibid., p. 252.

¹¹⁵ Ibid., p. 105ff, 115.

because this principle is widely, and in various legal systems, national, international, and supranational, regarded as an important element of the concretization of guarantees of human dignity.¹¹⁶ Human dignity, in turn, is a constitutive part of the whole architecture of human rights. But not only this. Human rights as such are seen as products of the mental gizmo, useful as rhetorical devices, and exploitable for good causes but without any claim to rationality as such and often quite harmful in their effects.¹¹⁷ Instead, for truly rational moral thinking one has to resort to utilitarianism. Utilitarianism is slow thinking that should govern human moral reasoning in the last instance.¹¹⁸

What is the evidence for the mental gizmo thesis? The familiar trolley problems are the starting point of the analysis.¹¹⁹ It is argued that a proper analysis of these cases shows that in cases where deontological judgments seem to be at play, because the participants judge that it is not permitted to sacrifice the life of one person to save five (footbridge case),¹²⁰ hard-wired emotional reactions are in fact determining the judgment.¹²¹ Decisive reasons for this analysis are neuro-imaging studies. These studies, it is argued, show that when deciding these cases, the ventromedial prefrontal cortex (VMPFC) is active, associated with the production of emotion.¹²² This is so because these cases involve the agent directly; they are “personal” and thus trigger emotional reactions.¹²³

In other cases that are “impersonal” and where the judgment is different and participants regard it as permitted that one person die and five persons be saved

¹¹⁶ Cf. M. Mahlmann, Human Dignity and Autonomy in Modern Constitutional Orders, in: M. Rosenfeld/A. Sajo (eds.), *The Oxford Handbook of Comparative Constitutional Law*, 2012, p. 370ff.

¹¹⁷ J. Greene, *Moral Tribes*, 2013, p. 302ff.

¹¹⁸ Ibid., p. 290ff. Utilitarianism is called “deep pragmatism.” Greene sums this thesis up: “The Central Tension principle: Characteristically deontological judgments are preferentially supported by automatic emotional responses, while characteristically consequentialist judgments are preferentially supported by conscious reasoning and allied processes of cognitive control,” idem., Why Cognitive (Neuro)Science Matters for Ethics, *Ethics* 124 (2014), p. 695ff, 699. For endorsement, cf., e.g., P. Singer, Ethics and Institutions, *The Journal of Ethics* 9 (2005), p. 331ff.

¹¹⁹ J. Greene, *Moral Tribes*, 2013, p. 105ff.

¹²⁰ In the footbridge case, a person is thrown down onto a track from a bridge to stop a runaway trolley to save people working on the track.

¹²¹ J. Greene, *Moral Tribes*, 2013, p. 119ff; idem., Why Cognitive (Neuro)Science Matters for Ethics, *Ethics* 124 (2014), p. 695ff, 698, states that mechanisms of fast thinking do not need to be “hard wired”. In his discussion, the gizmo appears, however, throughout to be “hard wired”, a given of human nature, shared by Kant, Rawls, and us.

¹²² J. Greene, *Moral Tribes*, 2013, p. 121ff; J. Greene/R. B. Sommerville/L. E. Nystrom/J. M. Darley/J. D. Cohen, An fMRI investigation of emotional engagement in moral judgement, *Science* 293/5537 (2001), p. 2105ff.

¹²³ J. Greene, *Moral Tribes*, 2013, p. 121ff.

(bystander case),¹²⁴ in contrast the dorsolateral prefrontal cortex (DLPFC), the center of cognitive control in the brain, is active, showing that this utilitarian judgment is a rational, not an emotional judgment.¹²⁵ Further evidence are studies that show, it is argued, that if the VMPFC is damaged, utilitarian judgments are made by the participants in both the bystander and footbridge cases.¹²⁶

Deontological arguments are from this perspective nothing but the post hoc rationalization of such hard-wired emotional reactions.¹²⁷ This is not only the “secret joke of Kant’s soul”¹²⁸ but, one might add, the secret joke of the soul of a great many thinkers on moral issues from antiquity to contemporary endeavors in this respect. Thus, say, Plato’s dialogues, the *Critique of Practical Reason*, or *A Theory of Justice* are all exercises in self-delusion of their authors and recipients: They miss the central point that Plato’s defense of a nonconsequentialist concept of justice, the formulation of the categorical imperative in its formal and material versions or the nonutilitarian principles of justice of John Rawls are ultimately the expression of the secret working of the emotional mental gizmo rationalized post hoc and writ large.¹²⁹ Human rights are part of these post hoc rationalizations: “‘Rights’ are nothing short of brilliant. They allow us to rationalize our gut feelings without doing any additional work.”¹³⁰ The many people concerned with human rights such as lawyers, judges, activists, politicians, and – most importantly – people claiming human rights, fighting for their respect, hoping, sometimes desperately, for their protection are all under the spell of a “moral illusion.” Given the importance of human rights in practice, this “moral illusion” has massive consequences that belittle any practical effects that other elements of a skewed rationality like framing effects may possibly have. These consequences are quite harmful:

¹²⁴ In the bystander case, the bystander can turn a switch so that a runaway trolley is redirected with the consequence that it does not kill five persons on one track but one person on another track.

¹²⁵ J. Greene, *Moral Tribes*, 2013, p. 120.

¹²⁶ Ibid., p. 124ff. For more studies taken as support for this thesis, idem., Why Cognitive (Neuro) Science Matters for Ethics, *Ethics* 124 (2014), p. 695ff, 700ff.

¹²⁷ Idem., *Moral Tribes*, 2013, p. 298ff, 300: “The moral equivalent of confabulation is *rationalization*. The confabulator perceives himself doing something and makes up a rational sounding story about what he’s doing and why. The moral rationalizer *feels* a certain way about a moral issue and then makes up a rational-sounding justification for that feeling” (emphasis in the original).

¹²⁸ Ibid., p. 301, quoting Nietzsche: “In other words, Kant has the same automatic settings as is surrounding tribespeople. But Kant, unlike them, felt the need to provide esoteric justifications for their ‘popular’ prejudices.”

¹²⁹ On J. Rawls, *A Theory of Justice*, 1971, as being another product of rationalization of the working of the gizmo, J. Greene, *Moral Tribes*, 2013, p. 333.

¹³⁰ J. Greene, *Moral Tribes*, 2013, p. 301ff, 302.

“Rationalization is the great enemy of moral progress, and thus of deep pragmatism.”¹³¹

The mental gizmo thesis is part of a wider current in contemporary neuroscience that one may call neuroscientific emotivism, defending the idea of traditional emotivism that human morality is nothing but the expression of certain emotions of appraisal and disgust with the means of cognitive psychology.¹³² It is a critique of human rights of a very radical kind. Its punch line is not to deny the cognitive reality and impact of deontological judgments and the idea of human rights, but to radically reinterpret their status and meaning: They are not a manifestation of practical reason but on the contrary part of the causes of human moral irrationality. This irrationality has such far-reaching detrimental consequences that it has to be overcome for the sake of the survival of the species. Only utilitarianism, slow thinking, can, it is maintained, solve the great problems of humanity and transcend the parochial moralities of the human tribes created by the mental gizmo,¹³³ “to free philosophers from the ups and downs of their automatic settings,”¹³⁴ and all the rest of us, of course, who are also often suffering from the “moral illusion” of human rights.

How to answer this challenge that is interesting and presumably paradigmatic for quite a few discussions to come? This is the next question to be addressed.

3 *The Mental Gizmo Thesis Reconsidered*

A fundamental problem is that the analysis of the trolley problems underlying the mental gizmo thesis is deficient.¹³⁵ The cases that are taken to prove the working of emotional gut reactions (footbridge) are in fact showing something quite different, namely the relevance of the means / side effect distinction for the explanation of the empirical patterns of moral evaluation observed, and

¹³¹ Ibid., p. 301.

¹³² Cf., e.g., J. Haidt, *The Emotional Dog and Its Rational Tail*, *Psychological Review* 108 (2001), p. 814ff; idem., *The Righteous Mind*, 2012; S. Nichols, *Sentimental Rules*, 2004. There are some qualifications, e.g., in Haidt's work the possibility of social influences on what he regards as moral evaluations.

¹³³ J. Greene, *Moral Tribes*, 2013, p. 289ff.

¹³⁴ Idem., *Why Cognitive (Neuro)Science Matters for Ethics*, *Ethics* 124 (2014), p. 695ff, 720.

¹³⁵ The usual reference for the origin of this problem is to P. Foot, *The Problem of Abortion and the Doctrine of the Double Effect*, *Oxford Review* 5 (1967), p. 5ff; J. J. Thomson, *The trolley problem*, *Yale Law Journal* 94 (1976), p. 1395ff. In fact, the problem has been formulated before, cf. H. Welzel, *Zum Notstandsproblem*, *ZStW* 1951, p. 47ff, 51 (with trains, not trams as in Foot). It is interesting that one finds almost no reference in current international debates to this earlier paper, though Welzel was a well-known criminal lawyer and the article has been a standard piece in German-language criminal law discussions ever since it was published. Whether Foot was familiar with it would be interesting to know.

more precisely, the prohibition of the instrumentalization of a human being.¹³⁶ A proper analysis of the trolley cases therefore in fact vindicates the relevance of the principle that Kant framed in the famous formula of the principle of humanity and thus a central element of the idea of human rights. In addition, it is far too rash to conclude that in other cases, where choosing the death of one person is taken to be permissible if the alternative is the death of five (or more victims, as in Welzel's initial train case), a utilitarian reasoning is at play. To think that it is permissible to choose the lesser of two unavoidable evils, though perhaps feeling at the same time that this is a tragic choice, is one thing; to endorse utilitarianism in the sense that it is always permitted to just count lives quite another. The analysis of the trolley cases must be much more sophisticated and transcend such simple dichotomies.¹³⁷

¹³⁶ The most advanced analysis of the trolley problem is provided by J. Mikhail, *Elements of Moral Cognition*, 2011, including the introduction of formal modes of representation of structures of human actions evaluated like act trees, cf. *ibid.*, p. 118ff. Greene discusses Mikhail in some detail, cf. J. Greene, *Moral Tribes*, 2013, p. 230ff. As Mikhail correctly argues, the core of the footbridge scenario is the use of the patient as a means. More precisely, in Mikhail's analysis, a battery (not the death) is a means to stop the trolley, not just the foreseen side-effect of an action to save five, *idem.*, *Elements of Moral Cognition*, 2011, p. 123ff. One may question whether the means to stop the trolley is just a battery or the death of the person, and this may be a significant difference. The means/ends-distinction is, however, in any case of crucial importance. An alternative explanation relying on the personal character of the action (pushing the patient over the bridge) is unconvincing, given scenarios that remove this personal element, cf. *ibid.*, p. 109 (drop man) and *passim*; *idem.*, *Moral Grammar and Human Rights: Some Reflections on Cognitive Science and Enlightenment Rationalism*, in: R. Goodman/D. Jinks/A. K. Woods (eds.), *Understanding Social Action, Promoting Human Rights*, 2012, p. 160ff, 183. Empirical evidence quoted by J. Greene, *Moral Tribes*, 2013, p. 215ff is inconclusive, given empirical evidence that on the contrary buttresses the relevance of means / side effect distinction, J. Mikhail, *Elements of Moral Cognition*, 2011, p. 319ff and other studies with quite different results, cf. below Fn. 147. The "loop case" does not put these findings into question. On this case adduced by J. Greene, *Moral Tribes*, 2013, p. 220ff as a counter example, J. Mikhail, *Elements of Moral Cognition*, 2011, p. 336ff, 359. In the latter case, the issue of what counts as the origin of data, i.e., the problem of the criteria for the selection of judgments taken as evidence, is of central importance, because some scenarios can be so complicated that their moral point becomes obscure; see the remarks on "considered judgments" below. The "modular myopia hypothesis" that J. Greene, *Moral Tribes*, 2013, p. 224ff, formulates as an explanation and according to which the emotional (deontological) cognitive subsystem is blind to harmful side effects is therefore not convincing.

¹³⁷ Cf. for an attempt to move forward in this respect J. Mikhail, *Elements of Moral Cognition*, 2011; on alternative explanations for the reactions to trolley cases, cf. G. Kahane/N. Shackel, Methodological Issues in the Neuroscience of Moral Judgment, *Mind & Language* 25 (2010), p. 561ff. As E. Zamir, *Law, Psychology, and Morality*, 2015, p. 188, has pointed out, given sufficiently large "net good outcomes", subjects may reason as if they were consequentialists. One important question in this context is, whether there is a significant moral difference between cases where harm is inflicted on others and which are widely regarded as (morally and legally) justified and such cases as the bystander scenario in the trolley case. The

Another problem of the mental gizmo thesis is that it is self-refuting. The reasons for this are as follows. Utilitarianism in all its classical and contemporary rule-based or action-based variants is guided by the principle of utility: Normatively justified is any rule or action that creates the greatest happiness of the greatest number.¹³⁸ The foundation of this principle is the idea that the happiness of any person counts equally. This is the egalitarian bite of utilitarianism that explains its persistent attraction and is the core of what is truly admirable in this line of thought.¹³⁹

This maxim rests on two pillars: The equality of persons and the prescriptive principle that equal persons ought to be treated equally. Only given these two propositions does it follow that everyone's happiness shall count equally, as presupposed by the principle of utility.

Consequently, the obligation to respect the equality of equal persons by counting their happiness equally in the utilitarian calculus is not a consequence, but a precondition of utilitarianism. Because that is so, the obligation to respect the equality of persons is not and cannot be derived from the application of the principle of utility: Being its foundation, the obligation to treat equals equally cannot be the consequence of that latter principle. The obligation to respect the equality of persons is rather a principle that is the nonconsequentialist normative precondition of consequentialism. There is thus a deontological remnant at the core of utilitarianism, because the principle of obligatory equal treatment of equals (the reason for being obliged to value the happiness of everybody equally as presupposed by the principle of utility) is the foundation of any utilitarian argument.¹⁴⁰

perception that the latter case involves tragic choices implies that there is. S. Nichols/R. Mallon, *Moral Dilemmas and Moral Rules*, *Cognition* 100 (2006), p. 530ff, distinguish between broken (apparently explicit) rules and permissibility all-things-considered. This distinction raises interesting issues that are important. The issue discussed here as an incident of tragic choices arises, however, exactly in the case where an action is regarded as permissible all-things-considered and consequently has to be accounted for.

¹³⁸ Cf. J. Bentham, *An Introduction to the Principles of Morals and Legislation*, I, note 1, July 1822, ed. J. H. Burns/H. L. A. Hart, 1996.

¹³⁹ Cf. J. Bentham, *An Introduction to the Principles of Morals and Legislation*, I, 13 note d, ed. J. H. Burns/H. L. A. Hart, 1996; J. S. Mill, *Utilitarianism*, in: J. Gray (ed.), *John Stuart Mill, On Liberty and other Essays*, 1991, V, p. 200 and 198, on the principle of equality: "It is involved in the very meaning of Utility, or the Greatest Happiness Principle. That principle is a mere form of words without rational signification, unless one person's happiness, supposed equal in degree (with proper allowance made for kind), is counted exactly as much as another's."

¹⁴⁰ Greene realizes that the foundation of utilitarianism is such a principle of equality: idem., *Moral Tribes* 2013, pp. 163, 170, "The second utilitarian ingredient is impartiality, the universal essence of morality, that's distilled in the Golden Rule. Having added this second

This analysis of utilitarianism leads to the central catch of the gizmo thesis. The catch is this: Either it is true that utilitarianism is slow thinking. As utilitarianism presupposes deontological principles of equality, it follows that such deontological principles are in that case slow thinking, too, because these deontological principles of equality are the normative core of what is regarded as slow thinking. Or it is true that such deontological principles are fast thinking. As utilitarianism presupposes these deontological principles of equality, which is fast thinking, utilitarianism must be fast thinking, too. In either case, the gizmo thesis is refuted by internal contradictions.¹⁴¹

ingredient, we can summarize utilitarianism thus: Happiness is what matters, and everyone's happiness counts the same."

¹⁴¹ Greene does not account for the justification of the foundational principle of utilitarianism that he identified (impartiality, the Golden Rule, cf. *ibid.*, pp. 163, 170, Fn. 140) and draws no consequences from this, despite this evidently putting his analysis in question: What is "impartiality" or the "Golden Rule"? Slow thinking? Why is it not just another of those principles of ethics that he derides as dangerous rationalizations of gut reactions? What is the difference in this respect between "impartiality" and the "Golden Rule" and, say, the categorical imperative or Rawls' principles of justice? This question is to be asked not least because the point of the categorical imperative or Rawls' principles of justice is exactly to reach "impartiality" by universalization or by deliberation behind a "veil of ignorance". The "Golden Rule" is certainly related to one of the central ideas of the categorical imperative as well, namely the idea of universalization. Greene's argumentation is thus circular: Principles like the categorical imperative are criticized as dangerous rationalizations of emotional gut reactions on the basis of principles like the Golden Rule that are in fact similar to the very principles criticized.

Idem., Why Cognitive (Neuro)Science Matters for Ethics, *Ethics* 124 (2014), p. 695ff, 717, writes that his argumentation "favors consequentialist approaches to moral problem solving, ones aimed solely at promoting good consequences, rather than deontological approaches aimed at figuring out who has which rights and duties, where these are regarded as constraints on the promotion of good consequences." One can state the critique developed here in the terms of constraining rights and duties, too: The principle of equal treatment ("impartiality") at the base of utilitarianism implies that every human being has a *right* that her happiness should count equally and that the others have a *duty* to count her happiness equally. Only given these normative constraints is the application of the principle of utility, for utilitarianism, legitimate. The doctrine of rights and duties is thus *criticized*, using a doctrine that relies in turn on very important rights and duties. Again, the circle is complete. Greene even states that utilitarianism presumably rests on an "affectively based evaluative premise," *ibid.*, p. 724. That this "affectively based evaluative premise" is supposed to be a high level intuition does not change the fact that – by relying on such an *affectively* based premise to substitute emotional gut reactions – the theory has become quite visibly inconsistent. It is a useful exercise to reconsider on the basis of this observation the meaning of the studies (interesting as they are) for Greene's argument, which he lists in *ibid.*, p. 701ff. Another example of these kinds of contradictions is the following statement, referring to emotional and cognitive neural structures: "It seems that healthy humans engage both responses and that there is a higher-order evaluation process that depends on the ventromedial prefrontal cortex, a structure that across domains attaches emotional

Thinking about these kinds of claims of certain psychological theories one should keep one thing in mind: Neuroimaging studies have provided in recent years many fascinating results. There are, however, as of yet limits to these insights, too. These limits are not only the product of the constraints imposed and problems implied in the methods of neuroimaging, for instance, because of standard questions like the level of spatial and temporal resolution of the methods applied, “voodoo correlations”,¹⁴² circular analysis,¹⁴³ effects of the statistical “smoothing” of results and so on.¹⁴⁴ It is, for instance, still an unsettled issue how complex mental phenomena are realized in the human brain. One question in this context is whether the long dominant focus on the localization of function in the brain is fruitful or whether the research in patterns of activation is not more promising.¹⁴⁵ This has consequences for the interpretation of patterns of neuronal activity that are the basis for the

weight to decision variables. In other words, the brain seems to make both types of judgment (deontological and consequentialist) and thus makes a higher order judgment about which lower-order judgment to trust, which may be viewed as a kind of wisdom (reflecting virtue or good character),” J. Greene/F. Rossi/J. Tasioulas/K. Brent Venable/B. Williams, Embedding Ethical Principle in Collective Decision Support Systems, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, 2016, p. 4147ff. It is noteworthy that the last instance arbiter of human moral judgment is in this passage a brain region that is associated with emotional, deontological judgement, though here couched in the terms of some kind of virtue ethics. It is unclear how this is reconcilable with Greene’s theses on fast and slow moral thinking.

These problems affect other studies, too, e.g., on the effects of variation in the oxytocin receptor gene on moral judgment, that build on these assumptions, cf. R. Bernhard/J. Chaponis/R. Siburian/P. Gallagher/K. Ransohoff/D. Wikler/R. Perlis/J. Greene, Variation in the oxytocin receptor gene (OXTR) is associated with differences in moral judgment, *Social Cognitive and Affective Neuroscience* 11 (2016), p. 1872ff.

¹⁴² Cf. E. Vul/C. Harris/P. Winkielman/H. Pashler, Puzzling high correlations in fMRI studies of emotion, personality, and social cognition, *Perspectives on Psychological Science* 4 (2009), p. 274ff.

¹⁴³ Cf. N. Kriegeskorte/W. K. Simmons/P. Bellgowan/C. I. Baker, Circular analysis in systems neuroscience: the dangers of double dipping, *Nature Neuroscience* 12 (2009), p. 535ff.

¹⁴⁴ Cf., e.g., R. A. Poldrack, The future of MRI in cognitive neuroscience, *Neuroimage* 62 (2012), p. 1216ff on some statistical problems.

¹⁴⁵ Cf., e.g., *ibid.*, 1217f on the move away from “blobology” to pattern analysis: “the goal of finding blobs in a specific region can drive researchers into analytic gymnastics in order to find a significant blob to report. However, for the last few years the most interesting and novel research has focused on understanding patterns of activation rather than localized blobs. The appreciation of patterns is happening at multiple scales. At the systems (whole-brain) scale, the modelling of connectivity and its relation to behaviour continue to grow ... I think the jury is still out on how well fMRI can ever characterize neuronal connectivity; as we outlined in Ramsey et al. (2010), there are a number of fundamental challenges in using fMRI to characterize causal interaction between brain regions.” For a similar assessment (from phrenology to network theories) cf. L. Jäncke, *Kognitive Neurowissenschaften*, 2013, p. 71ff.

hypotheses of neuroimaging studies because it is far from clear what certain observed neuronal activity really means. This is not least due to a classical problem of neuroimaging studies: the problem of reverse inference. From the fact that a brain region is active when performing a certain task one cannot conclude that whenever this brain region is active, this cognitive task is performed. The reason is that a brain region may perform many tasks, interacting for instance with other brain areas. Reverse inference can produce research hypotheses, but no conclusive evidence, for how a mental function maps onto the brain.¹⁴⁶ That a brain region, say the DLPFC, is active when performing certain tasks of cognitive control does thus not entail that whenever the DLPFC is active, tasks of cognitive control are performed. The same is true of the VMPFC and emotional reactions. Consequently, the results of neuroimaging studies such as the ones referred to, if they stand the test of further research, are in no way proof that deontological morality is the expression of emotion. In addition, there are many competing empirical findings.¹⁴⁷

¹⁴⁶ Cf. R. A. Poldrack, Can cognitive processes be inferred from neuroimaging data?, *Trends in Cognitive Science* 10 (2006), p. 59ff; idem., Inferring Mental States from Neuroimaging Data: From Reverse Inference to Large-Scale Decoding, *Neuron* 72 (2011), p. 692ff; idem., The future of fMRI in cognitive neuroscience, *Neuroimage* 62 (2012), p. 1216ff, 1216, 1218f, on the (difficult) task of finding “a region that is engaged *selectively*, such that activation of the region is actually predictive of the mental process” (emphasis in the original), as a precondition for overcoming the problems of reverse inference.

¹⁴⁷ One should not overestimate the state of insight achieved by experimental work. To take an example, there are studies that suggest that VMPFC patients are more vindictive in ultimatum games than normal subjects (cf. M. Koenigs/D. Tranel, Irrational Economic Decision-Making after Ventromedial Prefrontal Damage: Evidence from the Ultimatum Game, *The Journal of Neuroscience* 27 (2007), p. 951ff), which seems to imply a less “utilitarian” and more “deontological”, fairness-oriented outlook, while the same brain defect is used as an argument for the thesis that “deontological” judgments are emotional reactions emanating from the VMPFC; see above. This is not really convincing as: “Such patients exhibit both an abnormal utilitarian *and* an abnormal deontological tendency!,” G. Kahane/N. Shackel, Methodological Issues in the Neuroscience of Moral Judgment, *Mind & Language* 25 (2010), p. 561ff, 573 (emphasis in the original). On the same problem, cf. A. A. Duke/L. Bègue, The Drunk Utilitarian: Blood Alcohol Concentration Predicts Utilitarian Responses in Moral Dilemmas, *Cognition* 134 (2015), p. 121ff, 124: “Alcohol intoxication is associated with increased emotional reactivity and selective attention towards emotional cues, which according to Greene’s dual process conceptualisation, should lead to increased deontological (non-utilitarian) inclinations, the opposite of what was observed here.” On another study with the result that “there is little relation between sacrificial judgments in the hypothetical dilemmas that dominate research, and a genuine utilitarian approach to ethics,” G. Kahane/J. A. C. Everett/B. D. Earp/M. Farias/J. Savulescu, Utilitarian judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good, *Cognition* 134 (2015), p. 193ff, 193. Related is the debate of J. Moll/R. de Oliveira-Souza, Moral Judgments, Emotions and the Utilitarian Brain, *Trends in Cognitive Science* 11 (2007), p. 310ff; J. Greene, Why Are VMPFC Patients More

Consequently, in the light of what has been said there is ample reason to reinterpret such findings and what they tell us about the working of the mind, considering closely the analysis above and plausible theories of moral judgment.

In this context, it should be underlined that nobody doubts that emotions are a central part of moral evaluation. There may even be emotions that are “geological upheavals” of moral thought.¹⁴⁸ Nor is the importance denied that certain emotions have for the design of law and not least for the skeptical project of constitutionalism.¹⁴⁹ The question is, however, whether such emotions constitute moral (deontological) evaluation, as emotivists contend, or play different roles. Again, the theoretical power of imagination seems in emotivist accounts much too limited to explain the complexity of the human moral world.¹⁵⁰ One important example is the analytical failure to distinguish emotions that are the *consequences* of moral judgment from emotions that *constitute* moral judgment. Consider the case of outrage after witnessing a grave injustice. The cognition of injustice is here the precondition of and thus not identical to the feeling of moral outrage. Another central function of emotion is to fathom what an action means for the patient of the action. It is a piece of moral heuristics: Without empathy for the victims of racial discrimination, for instance, the emotional understanding of how it feels to be degraded and humiliated, nobody will be able to evaluate its injustice properly.

In addition, as the sketchy analysis of the concept of a fundamental or human right has shown, the predication of rights is highly complex, in form as

Utilitarian? A Dual Process Theory of Moral Judgement Explains, *Trends in Cognitive Science* 11 (2007), p. 322ff; J. Moll/R. de Oliveira-Souza, Response to Greene: Moral Sentiments and Reason: Friends or Foes?, *Trends in Cognitive Science* 11 (2007), p. 323f. Another move is to reinterpret findings on the trolley problem in the framework of “intuitive/counterintuitive judgements,” cf. G. Kahane/K. Wicth/N. Shackel/M. Farias/J. Savulescu/I. Tracey, The Neural Basis of Intuitive and Counterintuitive Moral Judgment, *Social Cognitive and Affective Neuroscience* 7 (2012), p. 393ff; on some reinterpretation of the role of the VMPFC in moral decision making J. Greene, The Cognitive Neuroscience of Moral Judgement and Decision Making, in: M. Gazzaniga/G. R. Mangun (eds.), *The Cognitive Neurosciences V*, 2014, p. 1013ff, 1017ff. One can conclude from these debates that moral and legal theory is urgently needed to create a theoretical framework in which experimental findings can be more successfully designed and interpreted, including a much more fine-grained account of the structure and content of morality and the role of emotions as a precondition and consequence of moral judgment than is sometimes used in these experiment-based debates.

¹⁴⁸ To use the Proust metaphor of M. Nussbaum, *Upheavals of Thought*, 2001, p. 1.

¹⁴⁹ Cf. A. Sajó, *Constitutional Sentiments*, 2011, not least on fear.

¹⁵⁰ Cf. M. Mahlmann, Cognitive Science, Ethics and Law, *German Law Journal* 8 (2007), p. 586ff; M. S. Pardo/D. Patterson, *Minds, Brains, and the Law*, 2013, p. 58, on emotions accompanying moral judgments.

in substance. To equate such complex judgment with an emotional gut reaction does not seem to be a very promising approach to the matter.

All this is nothing but a reminder of the theory dependence of the interpretation of empirical data: Data only have meaning within a theoretical framework. Concretely, the worth of neuroimaging studies about the neurophysiological basis of moral judgment is dependent on the merits of the theoretical framework they are developed in. If this framework is deficient, the interpretation of the data will be insufficient, too.¹⁵¹

In addition, it seems to be a rather naïve idea to identify rationality with utilitarianism.¹⁵² The question is: Why should the scope of practical reason, to use a traditional term, not be wider? In the history of thought, it was the default assumption that human thought is not only made up of instrumental rationality of some sort but that there are other, qualitatively different yardsticks, most importantly of justice and moral goodness. This is the common denominator of much of the greatest thought on these matters. What is wrong with this? Why are deontological principles *a priori* not rational, or reasonable, if you prefer? What is intrinsically better about the principle of utility (forgetting for a moment its deontological foundations) in comparison with the prohibition of instrumentalization, the principle of the justice of equal treatment of equals or the obligations for care for others?¹⁵³ There is a strange impoverishment of the richness of human thought at play in such theories that does not foster the insights of practical philosophy and legal theory.

It is important to emphasize that the mental gizmo thesis is no necessary consequence of the dual process model of the mind as such. One can think that this model describes an important aspect of human cognition without finding the mental moral gizmo thesis convincing. Deontology could be part

¹⁵¹ “Armchair philosophers” are often criticized in current debates for their naïveté, sometimes with good reasons. One should, however, not overlook the deficits of some of the experimental work that could profit a great deal from more prior theoretical work (cf. above, Fn. 147).

¹⁵² J. Greene, *Why Cognitive (Neuro)Science Matters for Ethics*, *Ethics* 124 (2014), p. 695ff, 696, posits that slow thinking is “a general-purpose reasoning system, specialized for enabling behaviors that serve long[er] term goals.” This overlooks the insight from the theory of justice that equality as a normative principle is not the same as general rationality; cf. S. Gosepath, *Gleiche Gerechtigkeit*, 2004; S. Baker, *The Normative Insignificance of Neuroscience*, *Philosophy and Public Affairs* 37 (2009), p. 293ff, underlines correctly that a normative argument is needed to justify the conclusion that deontological judgments use morally irrelevant criteria, whereas utilitarianism does not. These normative criteria cannot be drawn from neuroscientific research as such, *ibid.*, p. 326.

¹⁵³ J. Greene, *Moral Tribes*, 2013, p. 136: “Reasoning, as applied to decision making, involves the conscious application of decision rules.” Why is according to this rather wide definition the principle of utility (or the principle of impartiality or the “Golden rule,” see above, Fn. 141) a candidate for reasoning but the categorical imperative or Rawls’ principles of justice are not?

of slow thinking; there is no a priori reason why this could not be the case. Consequently, not being convinced of the mental gizmo thesis does not entail anything about the worth of the dual process model of the mind. The mental gizmo thesis is just an implausible thesis developed within this model of the human mind.

Nor do neuroscientific or psychological approaches to questions of ethics and law wed one as such to a certain perspective on the cognitive origin of these orders. Nothing in the theory of mind, in neuroscientific research or psychology, forces one in particular to develop an impoverished account of human practical thought. The answer to psychological skepticism is consequently not to ignore neuroscience and psychology, or save oneself in normative theory where psychology, whatever it says, simply does not count, but to develop a substantial concept of human moral cognition as an element of a wider theory of human rights that is more plausible than its skeptical alternatives. Given the state of affairs, there is no theory of human rights without a plausible theory of human normative cognition. And, to be sure, there are various theoretical alternatives to make some progress in this respect.¹⁵⁴ One of them is the mentalist approach of the universal moral grammar model that has attracted some attention in recent years and deserves perhaps to be explored in even more depth.

4 *The Mentalist Theory of Ethics and Law*

a The Mentalist Framework

A mentalist model of moral cognition investigates the question whether there are identifiable generative principles of moral judgment specific to human moral cognition universal and uniform across the species.¹⁵⁵ The background

¹⁵⁴ Moral psychology is a very vivid field; cf. for an overview, e.g., the contributions in J. M. Doris (ed.), *The Moral Psychology Handbook*, 2010.

¹⁵⁵ Cf., e.g., N. Chomsky, *Language and Problems of Knowledge*, 1988, p. 152; M. Mahlmann/J. Mikhail: Cognitive Science, Ethics and Law, in: Z. Bankowski (ed.), *Epistemology and Ontology*, 2005, p. 95ff; J. Mikhail, *Elements of Moral Cognition*, 2011; J. Mikhail, Chomsky and Moral Philosophy, in: J. McGilvray (ed.), *The Cambridge Companion to Chomsky*, 2nd ed., 2017, p. 235ff; M. Mahlmann, *Rationalismus in der praktischen Theorie*, 2nd ed., 2009; idem., Cognitive Science, Ethics and Law, *German Law Journal* 8 (2007), p. 577ff; G. Harman, Using a Linguistic Analogy to Study Morality, in: W. Sinnott-Armstrong (ed.), *Moral Psychology*, Vol. 1, 2008, p. 345ff; E. Roedder/G. Harman, Linguistics and Moral Theory, in: J. M. Doris (ed.), *The Moral Psychology Handbook*, 2010, p. 273ff; R. Jackendoff, *Language, Consciousness, Culture*, 2007, p. 277ff; S. Dwyer, Moral Competence, in: K. Murusagi/R. Stainton (eds.), *Philosophy and Linguistics*, 1999, p. 169ff; M. D. Hauser, *Moral Minds*, 2006. For a recent critique, cf. M. S. Pardo/D. Patterson, *Minds, Brains, and*

approach has been highly successful in other domains of the theory of mind, for instance, prominently so in the study of language.¹⁵⁶ The principles are derived from the analysis of the practice and phenomenology of moral judgment. One complication of such a study, and a major one, is that moral judgments are intrinsically contested. A plausible way to proceed is to look at qualified moral judgments: These judgments have to be *considered judgments*, to borrow a useful term, in the sense that they are reflective, dispassionate judgments not skewed by interest, passion, error of facts, and so forth.¹⁵⁷ In the background is the distinction between competence and performance, the faculty to perform a certain cognitive task like moral judgment and the actual performance of this task.¹⁵⁸ There are only indirect conclusions possible from the performance to the competence of an agent because the performance is influenced by many other factors than just the structure of the competence. There is no doubt that loud techno music will influence the mathematical problem-solving capacity of people exposed to them. Nobody, however, entertains the idea that techno sound is of great relevance for the study of the cognitive apparatus enabling humans to do math. In the case of moral judgment, the final evaluation of an act can be biased, for example, by the interests of the evaluating person. Consequently, such influences must be factored out of the analysis to truly study the respective cognitive competence. This is central, and sometimes overlooked, in recent studies of moral psychology that apparently presume to study a human moral competence but are to a surprisingly large degree concerned with performance problems, such as the

the Law, 2013, p. 12ff, 63ff, especially because of the (externalist) thesis that there can be (on conceptual grounds) no unconscious rule following. On Wittgenstein's concept of rules underlying this argument, see M. Mahlmann, *Rationalismus in der praktischen Theorie*, 2nd ed., 2009, p. 121ff.

¹⁵⁶ Cf. N. Chomsky, *Aspects of a Theory of Syntax*, 1965; idem., *The Minimalist Program*, 1995; idem., *New Horizons in the Study of Language and Mind*, 2000.

¹⁵⁷ Cf. on this matter J. Rawls, *A Theory of Justice*, revised ed., 1999, p. 42: "Thus in deciding which of our judgments to take into account we may reasonably select some and exclude others. For example, we can discard those judgments made with hesitation, or in which we have little confidence. Similarly, those given when we are upset or frightened, or when we stand to gain one way or the other can be left aside. All these judgments are likely to be erroneous or to be influenced by an excessive attention to our own interests. Considered judgments are simply those rendered under conditions favorable to the exercise of the sense of justice, and therefore in circumstances where the more common excuses and explanations for making a mistake do not obtain"; J. Mikhail, *Elements of Moral Cognition*, 2011, p. 51ff.

¹⁵⁸ Cf. for the (crucial) competence/performance distinction N. Chomsky, *Aspects of the Theory of Syntax*, 1965, p. 3ff; M. Mahlmann, *Rationalismus in der praktischen Theorie*, 2nd ed., 2009, p. 73f; J. Mikhail, *Elements of Moral Cognition*, 2011, p. 51ff.

skewing of moral judgment by nonmoral factors from smells¹⁵⁹ to feelings of being controlled.¹⁶⁰ Another methodological step in order to deal with the contested nature of moral judgment is to look at highly idealized and often artificial cases that appear to be as little politically and culturally loaded as possible.¹⁶¹ To study the human moral faculty by studying opinions about abortion is, in contrast, a nonstarter.

b Properties of Moral Cognition

If one looks carefully at a phenomenology of morality, there are some important observations to be made.¹⁶² One is that there is something like a moral space at all. Humans operate naturally within a mental space that has a normative dimension. There is a specific mental domain of morality, a qualitatively distinguished element of conscious thought, an introspectively accessible, distinctive, intuited, subjectively experienced aspect of our mental life, a qualia, as it is often said, or – to use standard understandings of this term – a certain kind of phenomenal character of specific forms of experience. The availability of such a cognitive domain is not a self-evident fact but an empirical property of the human mind not common to all organisms. It does not concern a side issue but a property that defines nothing less than central elements of the identity of the human species: the moral dimension of their lives. Its existence consequently deserves close attention.

Another interesting observation concerns the fact that there is a highly and intricately qualified limited set of possible objects of moral evaluation. The kinds of possible moralities are already restricted by this set. The dropping of an apple from a tree in the hands of a hungry person is not a virtuous action of the tree. This is so because agency is a precondition of the moral evaluation of certain events in the world. If they are not attributable to agents, questions of moral evaluation do not arise.¹⁶³ To put your pen gracefully on a desk cannot be the possible object of moral evaluation, either, if there are not very particular circumstances, even if an agent performs this act. It is,

¹⁵⁹ S. Schnall/J. Haidt/G. L. Clore/A. H. Jordan, Disgust as Embodied Moral Judgement, *Personality and Social Psychology Bulletin* 34 (2008), p. 1096ff.

¹⁶⁰ Cf. K. Haley/D. Fessler, Nobody's Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game, *Evolution and Human Behaviour* 26 (2005), p. 245ff.

¹⁶¹ Cf. J. Mikhail, *Elements of Moral Cognition*, 2011, p. 56ff; M. Mahlmann, *Rationalismus in der praktischen Theorie*, 2nd ed., 2009, p. 107.

¹⁶² Cf. M. Mahlmann, *Rechtsphilosophie und Rechtstheorie*, 4th ed., 2017, p. 288ff.

¹⁶³ This does not mean that inanimate things have been taken as agents – cf. Xerxes' whipping of the sea to punish Poseidon.

however, a possible object of aesthetical evaluation – another distinctive element of human experience.¹⁶⁴ Another precondition of moral evaluation is – to put it very roughly – something like a volitionally controlled or controllable bodily action or omission of agents with consequences for the well-being of sentient beings, intentions of such actions or omissions, states of affairs resulting from such actions, or qualified emotions directed at the well-being of others.¹⁶⁵

c Principles of Morality

It is very important to distinguish motivational inclinations to behavior from moral evaluation. The inclination for an action, including what is called prosocial behavior, is one thing, the reflexive appraisal of this inclination with deontic dimensions quite another. Only the latter is the proper realm of morality and ethics.¹⁶⁶

If one turns to the content of morality and analyses carefully some qualified judgments of the kind described above, these judgments seem to be guided over a wide range of cases by principles of egalitarian justice and altruism.¹⁶⁷ This is not particularly surprising as it is a red thread in the history of ideas, too, that these principles form the core of morality. In recent years, some empirical work has been done that points in the same direction, though raising substantial theoretical problems.¹⁶⁸ The issues involved are complex, but one can

¹⁶⁴ On the classical distinction of the distinct and potentially contradictory moral and aesthetical evaluation, see I. Kant, *Kritik der Urteilskraft*, Akademie Ausgabe Bd. V, 1908/13, p. 204f.

¹⁶⁵ Intricate problems arise in this area. A matter of complex debate is, e.g., actions with effects for objects of art or actions affecting the environment. Especially the latter is of great practical concern. In both areas ethical principles matter. Another problem is not-other-regarding virtues. A fuller statement of the possible objects of moral evaluation would need to take account of these special cases, refining the basic principles stated here.

¹⁶⁶ This point is, e.g., relevant to the question of “animal morality.” A prosocial behavior of nonhuman animals does not as of itself constitute morality in the sense understood here. On the question of a possible continuum and differences between humans and (nonhuman) animals, e.g., J. Mikhail, Any Animal Whatever? Harmful Battery and Its Elements as Building Blocks of Moral Cognition, *Ethics* 124 (2014), p. 750ff; F. De Waal, *The Bonobo and the Atheist*, 2013; on some important differences between apes and humans M. Tomasello, *A Natural History of Human Morality*, 2016, p. 39ff.

¹⁶⁷ Cf. for some remarks M. Mählmann, *Rechtsphilosophie und Rechtstheorie*, 4th ed., 2017, p. 288ff.

¹⁶⁸ Research on justice in social psychology has a substantial tradition, encompassing research on relative deprivation, distributive justice, and the fairness of outcome distributions, procedural justice, and retributive justice; cf. T. R. Tyler/R. J. Boeckmann/H. J. Smith/Y. J. Huo, *Social Justice in a Diverse Society*, 1997, p. 11ff, for an overview of earlier research. A prominent example from the current debate is the extensive discussion of fairness or inequality aversion

tentatively identify the content of these principles in rough outline. For justice, the principle of equality is central. This is to a large degree the basis of discussions since antiquity. Perhaps the most important question in the theory of justice is what equality means exactly as a prescriptive concept. This is far from obvious, and influential voices have argued that it does not mean anything at all.¹⁶⁹

Such radical skepticism is, however, not entirely convincing because some identifiable elements seem to play an important role for the evaluation of the justness of actions. For an action to be just, the action has to satisfy arguably at least the following conditions: First, the standard of treatment applied has to be equal for all equal patients of the action. If one makes grading dependent on performance for some students and not for others, one

and altruism, e.g., in the case of contributions to social goods or (related) in the form of altruistic punishment. Importantly, strong reciprocity is at issue, that is, behavior that does not lead to individual economic benefit of the agent; cf. E. Fehr/U. Fischbacher, The nature of human altruism, *Nature* 425 (2003), p. 785ff. On cross-cultural research on the existence of such attitudes, cf., e.g., J. Henrich et al., “Economic man” in Cross-Cultural Perspective: Behavioural Experiments in 15 Small-Scale Societies, *Behavioural and Brain Sciences* 28 (2005), p. 795ff. On the development of such patterns of behavior in children, cf., e.g., K. R. Olson/E. S. Spelke, Foundations of Cooperation in Young Children, *Cognition* 108 (2008), p. 222ff (3, 5 year old); M. F. H. Schmidt/J. A. Sommerville (2011), Fairness Expectations and Altruistic Sharing in 15-Month-Old Human Infants, *PLoS ONE* 6 (2011), p. 1ff; E. Fehr/H. Bernhard/B. Rockenbach, Egalitarianism in Young Children, *Nature* 454 (2008), p. 1079ff; E. Fehr/D. Glätzle-Rützler/M. Sutter, The Development of Egalitarianism, Altruism, Spite and Parochialism in Childhood and Adolescence, *European Economic Review* 64 (2013), p. 369ff. On the effect of self-reflection in the framework of identity utility, cf. C. Engel/M. Kurschliigen, The Jurisdiction of the Man within – Introspection, Identity, and Cooperation in a Public Good Experiment, *Preprints of the MPI for Research on Public Goods* 1 (2015). This area of research provides many important findings. An important example of a theoretical problem of some of the studies is, however, that they are predominantly concerned with patterns of behavior (e.g., distributive acts, punishment, rewards), not with the reflective evaluation of actions which has deontic consequences. The latter is of central importance for human morality as moral judgment with such dimensions is the core of the matter, as already indicated. It is important in this respect to stay aware of a traditional insight of moral philosophy: There is no necessary connection between moral evaluation and moral behavior. An agent may very well perceive the morality or immorality of an action, but may fail to act accordingly, due to intervening interests, weakness of moral will, etc. If agents do not show altruistic or just behavior, this does not allow for direct conclusions as to the principles governing their moral judgment. Even if the behavior mirrors reflective evaluation, the question arises, whether or not considered judgments underlie this behavior. “Inequality aversions” or “other-regarding preferences” do not fully fathom the intricacies of moral judgments. Aversions and preferences describe inclinations to act. Moral judgment concerns something else, namely the evaluation of an intention or action (e.g., based on such inclinations) with deontic consequences, a moral ought. This “ought” is not a mere aversion or preference because of its obligatory force.

¹⁶⁹ Cf. H. Kelsen, Das Problem der Gerechtigkeit, in: idem, *Reine Rechtslehre*, 1960, p. 357ff.

grades unjustly. Second, if there is a criterion for distribution reasonably related to the sphere of justice¹⁷⁰ concerned (say performance for grading), a relation of equality has to be maintained between the value of the criterion and the kind of treatment – a good performance therefore deserves a good grade.¹⁷¹ If there is no such criterion, as a default rule, a numerically equal distribution among equals is just. If there is no particular reason to distribute a birthday cake differently, an equal distribution is a just distribution. This is not an entirely banal principle as it is central for the allocation of scarce goods in a society for cases in which a criterion of distribution underdetermines the distribution of these goods and of chances to acquire these goods. As far as restorative justice is concerned, the restituting act has to equal the object restored.¹⁷² Finally, any treatment has to be reconcilable with the basic equality of worth of human beings.

Whether there is genuinely other-regarding altruistic behavior or whether any action beneficial to others is ultimately motivated by some self-interest of the agent, though perhaps refined and hidden, is one of the traditional questions of practical philosophy. As in the case of justice, this is a huge debate, today enriched by interesting empirical work often connected with ideas stemming from evolutionary psychology.¹⁷³ In this context, it is important to distinguish the question whether people are in fact acting because of a genuinely altruistic motivation from the question whether genuine altruism is the precondition for *evaluating* something as morally good. There is not much reason to believe that people excel in altruistic behavior in their lives. This observation tells us, however, nothing about the principles that guide moral judgment, for instance, of the prevalent selfish behavior around us.¹⁷⁴

¹⁷⁰ To borrow a common term, M. Walzer, *The Spheres of Justice*, 1983.

¹⁷¹ It is a persistent problem to determine what criterion of distribution should count; cf., e.g., the example of three children and a flute taken from A. Sen, *The Idea of Justice*, 2009, p. 12f: Should a flute belong to its maker, the one who can play it best or the one that needs it the most for his well-being? The idea of proportional justice has been a centerpiece of the theory of justice since antiquity; cf. for a classic statement Aristotle, *Nicomachean Ethics*, 1129a ff.

¹⁷² Cf. M. Muhlmann, *Rechtsphilosophie und Rechtstheorie*, 4th ed., 2017, p. 335ff.

¹⁷³ Cf. for an overview S. Stich/J. M. Doris/E. Roedder, Altruism, in: J. M. Doris (ed.), *The Moral Psychology Handbook*, 2010, p. 147ff. As ultimate evolutionary mechanisms, kin selection, direct reciprocity, indirect reciprocity, network reciprocity, and group selection play a prominent role; cf., e.g., M. Nowak, Five Rules for the Evolution of Cooperation, *Science* 314 (2006), p. 1560ff.

¹⁷⁴ The incongruence of justified moral principles and behavior is not a new observation, cf. T. Nagel, *The Possibility of Altruism*, 1970, p. 146: “To say that altruism and morality are possible in virtue of something basic to human nature is not to say that men are basically good. Men are basically complicated; how good they are depends on whether certain conceptions and ways of thinking have achieved dominance, a dominance which is

Concerning these principles of evaluation, there are reasons to think that such genuinely altruistic motivation is in fact a core element of moral evaluation. More precisely, it seems plausible to assume that an action is morally good if it is performed with the direct intent, and not only the oblique intent, to foster the well-being of the patient. If this is so, it is irrelevant for the moral evaluation whether or not the fostering of the interests of the agent is – at the same time – directly intended or a foreseen (obliquely intended) consequence of the action and forms a second reason for action in a bundle of motives. The direct intention to foster the well-being of the patient of the action appears to be a necessary condition of morally good action.¹⁷⁵

To illustrate the meaning of this principle it is useful to look at one of the most refined versions of ethical egoism. This form of egoism holds that altruistic behavior is ultimately motivated by the desire to experience the satisfaction of having acted in a morally appropriate manner. There is an important point in this argument, namely that moral action indeed does provide some particular form of satisfaction for the agent and that agents are certainly often aware of this. In addition, immoral acting can have unpleasant effects, too, such as shame. These observations, however, do not settle the issue. Consider the following case: Person A helps another person with the thought “I do not care for this person and her well-being at all (what a silly person she in fact is!), it just happens (unfortunately) that I have to do something for her to reap the sweet fruit I really desire, namely to feel the satisfaction of being a truly nice person!” Is this really a morally laudable deed? If there are some doubts about the moral praiseworthiness of an action with such an intention, it seems to confirm the analysis above. This is so because the agent has only an oblique intention to help the other person and not the direct intention to be beneficial to her: Her direct intention is to satisfy one of her own personal desires and helping the other person is only an (perhaps even unwelcome) means to achieve that end.

Another point is perhaps worth noticing. Justice seems to be something like a limiting condition of morally good action: There is no morally good intention that violates principles of justice. If a person helps for instance three out of four people in need, not because she cannot help all but just because she feels like excluding one due to a whim, this is not a morally good action, despite her direct intention to help the other three, because it violated principles of equal treatment.

precarious in any case. The manner in which human beings have conducted themselves so far does not encourage optimism about the moral future of the species.”

¹⁷⁵ Cf. M. Mahlmann, *Rechtsphilosophie und Rechtstheorie*, 4th ed., 2017, p. 289ff. It is assumed that the evaluation of the action is dependent on the nature of the underlying intention.

Such tentatively outlined principles, which are obviously in need of much refinement, are abstract but not without meaningful content, as can be exemplified by the justification of human rights:¹⁷⁶ The principles underlying the attribution of rights to persons have to be equal for all potential bearers. It would be, for example, unjust to let some people enjoy fundamental rights because of their personhood, and deny them to others because for them skin colour (and not their personhood) is taken to be relevant. In addition, the reasonable – and more precisely only reasonable – criterion for the attribution of rights is the humanity of any person. As a just system of rights has to preserve a relation of equality between the value of this criterion and the distribution of rights, and as all humans are equal in their humanity, only a system of equal rights is consequently a just system of rights. To foster the enjoyment of rights is morally good and thus, as has been said before, given the importance of the goods that rights protect and of the rights themselves, the promotion of rights is a *prima facie* obligation of human solidarity.

d Volitional and Emotional Consequences of Moral Judgment

Another important point for the topic pursued is that human moral judgment has volitional consequences: A moral evaluation does not yield information about a fact of the world like a descriptive proposition; it has prescriptive content.¹⁷⁷ To say “X is just” is different from “X is blue” and an important element of the difference is that a moral judgment has volitional consequences if there is a possibility to act.¹⁷⁸ A normative statement is about an obligation, a permission, or prescription; in short, about a moral ought. To be sure, human motivation encompasses many other inclinations that have great power and have nothing to do with moral considerations. Human history is to

¹⁷⁶ Another example that these principles are not meaningless is that, e.g. Rawls’ principles of justice can be derived from them: The first principle of universal freedom and the principle of equal access to office are principles of equally distributed goods, freedom, and offices, respectively. The difference principle is a prudential modification of an egalitarian distribution of material goods in a society.

¹⁷⁷ M. Mahlmann, *Ethics, Law and the Challenge of Cognitive Science*, *German Law Journal* 8 (2007), p. 599ff.

¹⁷⁸ Cf. for a concise statement R. Price, *A Review of the Principal Questions in Morals*, D. D. Raphael (ed.), 1758, p. 186: “When we are conscious that an action is *fit* to be done, or that it *ought* to be done, it is not conceivable that we can remain *uninfluenced*, or want a *motive* to action” (emphasis in the original). On the background debate of motivational externalists and internalists, e.g., R. M. Hare, *The Language of Morals*, 1952, p. 20, 30, 169, 197; idem., *Moral Thinking*, 1982, p. 23; D. O. Brink, *Moral Realism and the Foundations of Ethics*, 1989, p. 39; G. Harman, *Explaining Value*, 2000, p. 30; P. Foot, *Virtues and Vices*, 1978, p. 148; J. L. Mackie, *Ethics: Inventing Right and Wrong*, 1977, p. 40.

a large extent the history of greed and the pursuit of power, not the history of moral niceties. The claim made is thus only that an element, and perhaps a precious element, of human moral motivation derives from moral insight.¹⁷⁹

The prescriptive content can – and that is important for the topic discussed – constitute a right.¹⁸⁰ If an act would be just, the agent has an obligation to act justly, and the patient has a right to that action: In the limited time available to comment and set things in intellectual order after a lamentably lousy plenary lecture, every discussant has the right to the same amount of time because this is a just distribution of this scarce good. The chair of the session has the obligation to ensure this fair distribution of time. The connection between obligation and right holds for an obligation stemming from a duty to benefit somebody, too, unless it is a supererogatory action: There is not only the obligation to pick up your phone to call an ambulance if somebody in front of you collapses, the person who has collapsed has a right that you do (at least) this. This is the reason why there is no theory of rights without a theory of justice and of justified obligation to others: This theory is the birthplace of rights.

e Questions of Meta-Ethics and Mentalism

Principles of justice and altruism that may guide the reflexive evaluation have cognitive content. Whether or not there is, for example, a relation of equality between patients of actions or between a criterion of distribution and the good distributed in the sense explained is not felt like coldness or perceived like roughness of a surface, but a predication, stemming from a complex structural analysis¹⁸¹ of the evaluated act predicating at the end a relation of equality, or its absence, and thus a judgment with cognitive content.

¹⁷⁹ Cf. M. Mahlmann, *Rationalismus in der praktischen Theorie*, 2009, p. 158ff; J. Mikhail, Moral Grammar and Human Rights: Some Reflections on Cognitive Science and Enlightenment Rationalism, in: R. Goodman/D. Jinks/A. K. Woods (eds.), *Understanding Social Action, Promoting Human Rights*, 2012, p. 169ff.

¹⁸⁰ Cf. above the analysis of rights and the connection of duties and (claim-)rights. On the relation of moral judgment and rights, cf. J. Mikhail, *Elements of Moral Cognition*, 2011, p. 295ff; idem., Moral Grammar and Human Rights: Some Reflections on Cognitive Science and Enlightenment Rationalism, in: R. Goodman/D. Jinks/A. K. Woods (eds.), *Understanding Social Action, Promoting Human Rights*, 2012, p. 160ff; M. Mahlmann, The Cognitive Foundations of Law, in: H. Rottluthner (ed.), *Foundations of Law*, 2005, p. 75ff; idem., *Elemente einer ethischen Grundrechtstheorie*, 2008, p. 517ff.

¹⁸¹ Cf., for an example of how complex such analysis is, the discussion of the trolley cases, J. Mikhail, *Elements of Moral Cognition*, 2011, p. 77ff.

Given the great variety of moral opinions today and in history, any theory of moral cognition has to formulate as part of its explanatory enterprise a theory of moral disagreement. This is yet another traditional and vast topic of practical philosophy. There is empirical work on this, too.¹⁸² It is sometimes argued that the mere fact of moral disagreement proves moral relativism.¹⁸³ In this respect, however, one should not draw conclusions too rashly. There may be ways to account for moral disagreement, even radical disagreement, with substantial explanatory power. A factor with considerable importance for explaining the existence of moral disagreement seems to be disagreement about the nonmoral preconditions of moral judgment. Others that rightly play a prominent role in this debate are the influence of interests or the impact of ideological constructions. Taking account of such factors may reduce the cases of real moral disagreement considerably. Take the (important) example of the rights of women: The denial of equal rights of women was partly based on wrong factual assumptions, for instance, on the idea that women lack the ability of autonomous self-determination or rationality and therefore have to be guided by men. Interests of men in comfortable structures of domination were an evident other factor. Ideological constructions, partly in a religious cloak, buttressed such social structures, too. Such factors continue to play an important political role today, despite the progress made. If such false assumptions about the nature of women or the power of interests and ideologies lose their influence, apparently irreconcilable moral disagreement can disappear quickly, and the equality of the rights of women may even appear across cultural borders as an evident truism (as it should).¹⁸⁴ There are therefore reasons to believe that under the surface of insurmountable moral disagreement a deep structure of common moral principles may be shared.¹⁸⁵

¹⁸² Cf. the attempts to explain different reactions, e.g., to insults by different “cultures of honours” in the US-American North and South and other issues, R. Nisbett/D. Cohen, *Culture of Honor: The Psychology of Violence in the South*, 1996; R. Nisbett, *The Geography of Thought: How Asians and Westerns Think differently . . . and Why*, 2003; J. Haidt, *The Righteous Mind*, 2012, p. 11ff. The disagreement can encompass the domain of morality as such, *ibid.*, p. 14ff.

¹⁸³ J. L. Mackie, *Ethics: Inventing Right and Wrong*, 1977, p. 36.

¹⁸⁴ Empirical studies like the ones referred in Fn. 182 are consequently far from conclusive. That there are many, often very different moral codes with significant practical importance, nobody doubts. That there were (and are), e.g., concepts of honor people feel strongly about that move others less or not at all is certainly true. But the importance of the possibility of criticism of such codes (say of the European duel) should not be underestimated, either. The conditions of the success of such criticism are a key to understand principles that underlie human moral deliberation and judgment that may finally lead to the abandonment of such practices.

¹⁸⁵ Cf. M. Mahlmann, Ethics, Law and the Challenge of Cognitive Science, *German Law Journal* 8 (2007), p. 593ff; J. Mikhail, Moral Grammar and Human Rights: Some Reflections on

As far as the ontology of morality is concerned, it is plausible to take moral cognition so conceived as being nonreferential: There are no objective moral facts in the world the correspondence with which is the truth condition of moral predicates.¹⁸⁶ This renders moral judgments not merely subjective. The reasons for this are internal mental yardsticks for justified propositions authenticating their truth as in other areas of thought.¹⁸⁷ Even a moral realist ontology ultimately relies on the assumed truth of such nonreferential statements: The moral realist thesis that a truth condition of moral judgments is that moral predicates correspond to objective moral facts in the world does not itself correspond to an objective epistemic fact in the world. Its truth depends thus on other sources of epistemic justification.¹⁸⁸

Cognitive Science and Enlightenment Rationalism, in: R. Goodman/D. Jinks/A. K. Woods (eds.), *Understanding Social Action, Promoting Human Rights*, 2012, p. 170ff.

¹⁸⁶ On this cf. J. Mikhail, *Elements of Moral Cognition*, 2011, p. 317; M. Mahlmann, Ethics, Law and the Challenge of Cognitive Science, *German Law Journal* 8 (2007), p. 580ff. For a defense of the view that there are, to the contrary, objective, irreducibly normative facts, cf., e.g., R. Shafer-Landau, *Moral Realism*, 2003; D. Enoch, *Taking Morality Seriously*, 2011. A nonreferential theory of ethics does not commit one to noncognitivism, desire-based ethics, expressivism, and the like, as explained in the text. The debate between moral realists and antirealists – as it stands today – does not exhaust the theoretical possibilities.

¹⁸⁷ That there are genuine normative reasons the truth of which does not depend on correspondence with entities that are part of the nonmental fabric of the world is defended from different points of view. Cf. C. C. Korsgaard, *The Sources of Normativity*, 1996, p. 108, 122ff, 165, arguing for a “reflective endorsement theory” that bases normativity on the self-endorsement of the humanity of the autonomous self; R. Dworkin, *Justice for Hedgehogs*, 2011, outlining an interpretative theory “all the way down”; D. Parfit, *On What Matters*, Vol. 2, 2011, arguing that there are “some irreducibly normative reason-involving truths”, which are “not about entities or properties that exist in some ontological sense”, *ibid.*, p. 618; T. M. Scanlon, *Being Realistic about Reasons*, 2014, developing a realistic “reasons fundamentalism.” For a critique of such theories, arguing in particular with the impossibility to distinguish on internal grounds true reasons and false moral propositions, rendering all argument indistinguishable from fictions, D. Enoch, *Taking Morality Seriously*, 2011, p. 121ff. The central point to counter this concern is that internal truth conditions of moral propositions are not at the whim of the agents. On the question of authentication of truth ultimately through foundational intuitions of truth, R. Jackendoff, *A User's Guide to Thought and Meaning*, 2012, p. 213ff, taking this as evidence that (in the terminology of the dual process model of the mind) system 2 (slow thinking) rides with the means of language on top of system 1 (fast thinking), without, however, making thinking irrational or emotional, because “it behoves us to show intuitive thinking more respect”, *ibid.* p. 215.

¹⁸⁸ Cf. on this argument recently, R. Dworkin, *Justice for Hedgehogs*, 2011, p. 76; T. M. Scanlon, *Being Realistic about Reasons*, 2014, p. 16, Fn. 1.

f The Development of Moral Cognition

One question to be explored in this context is the ontogenetic origin of moral cognition. The development of moral cognition is the object of landmark debates in moral psychology.¹⁸⁹ Given the analysis above, central for the question are – among other factors – the cognitive domain of morality, the restrictive principles that determine the possible objects of evaluation, the material principles of morality discussed, the prescriptive, volitional effects of moral judgments, the moral ought, including necessary connections between duties and rights or the emotional consequences of moral experience. Such phenomena could be constructed and acquired by secondary learning processes (instruction, repetition, etc.) as many normative principles are (e.g., the intricacies of Swiss law on unjust enrichment), or alternatively they could be at least in part the product of the unfolding of innate cognitive structures triggered by experience, as in the case of language. Whether or not this is the case is best answered by the poverty of stimulus argument: If the input by experience is not sufficient to generate a certain cognitive ability, at least some of the cognitive structures underlying this ability must be inborn.¹⁹⁰

It is important to emphasize that nothing in this approach denies the importance of social and cultural influences on ethics and law: The freedom of the press, for instance, as a legal norm and underlying principle of political morality presupposes the cultural achievement (a late achievement with plenty of preconditions) of the press and in addition the experience of its suppression even by democratic governments, among many other things. “The freedom of the press is protected” is certainly not an inborn principle of human moral cognition. The question, however, is: What kind of mind do you need to develop such an idea? What is special about the human mind that only humans and no other organism developed a concept like human rights? What are the cognitive preconditions for starting such a long cultural process, which leads after many thousands of years of human cultural and social

¹⁸⁹ E.g., of the work of J. Piaget, *Le jugement moral chez l'enfant*, 1932; or L. Kohlberg, *Essays on Moral Development*, Vol. I and II, 1981 and 1984.

¹⁹⁰ Cf. S. Laurence/E. Margolis, The Poverty of Stimulus Argument, *Brit. J. Phil. Sci.* 52 (2001), p. 217ff; M. Muhlmann, *Rationalismus in der praktischen Theorie*, 2nd ed., 2009, p. 74ff; J. Mikhail, *Elements of Moral Cognition*, 2011, p. 70ff. There are many interesting studies on the development of moral cognition; cf., e.g., the much discussed case of the moral-conventional distinction, L. P. Nucci/E. Turiel/G. Encarnacion-Gawrych, Children's Social Interaction and Social Concepts, *Journal of Cross Cultural Psychology* 14 (1983), p. 469ff, and the considerable research on this topic or the helper/hinderer experiment in J. K. Hamlin/K. Wynn/P. Bloom, Social Evaluation By Preverbal Infants, *Nature* 450 (2007), p. 557ff. For an example of a phylogenetic explanation driven by cooperation, M. Tomasello, *A Natural History of Human Morality*, 2016.

development to the idea that one can not only wish for or have an interest in but in fact enjoy a *right* to a free press? None? Just being smart? Or does one need certain specific conceptual tools with which to build a system of rights? Is there perhaps something to be discovered if one considers the moral cognitive space, the agency-dependency of moral judgment, the limited class of possible objects of moral evaluation, the principles of justice and altruism, and their foundational relation to rights? Is not the concept of “right” itself of great interest in this respect and what it entails, the intricate web of normative positions sketched above, the necessary connections between claims and duties, of privileges and negations of duties, the intentional content of deontic modes, the semantics of obligation, permission, and prediction, and the necessary volitional and emotional consequences of moral judgment and their implications for rights?

As illustrated by the mental gizmo thesis, the time has passed when the study of the structures of the mind seemed unimportant because it was taken for granted that only one kind of theory of mind is plausible, namely a theory that assumed that the only inborn property of the human mind is an unspecified learning device.¹⁹¹ The mental gizmo thesis is a substantial empirical thesis about the structure of the human mind, as any assumption about heuristics, framing effects or biases is. The latter motivates an enormous amount of research around the globe. These claims may be right or wrong, but they are certainly a serious scientific effort to be evaluated on their explanatory merits, and the same is true of the mentalist approach to ethics and law.

These remarks show that an empirically minded theory of moral psychology can be framed that does not take deontological principles as cognitive illusions but as part of the make-up of the human mind that may be the precondition of the cultural development of moral systems and the law. There is strong empirical evidence for the existence of a faculty of language with highly restrictive principles, in which natural languages unfold. There is, a great distance between the faculty of language to the verses of *King Lear*. The language faculty is, however, the precondition for the ability of humans to produce and enjoy something like *King Lear*. Similarly, it is perhaps a plausible idea that there is a great distance between the human moral faculty and the Universal Declaration of Human Rights, or other concrete, historically shaped catalogues of rights in constitutions and international bills of rights. But the moral faculty could turn out – as the faculty of language is for the verse of *King Lear* – to be the cognitive precondition for the possibility

¹⁹¹ Cf. the remarks in N. Chomsky, *Aspects of Theory of Syntax*, 1965, p. 47ff.

ultimately to frame something like the Universal Declaration of Human Rights and the aspirations it implies.

5 Explanatory and Normative Theory

Let us assume that a mentalist account of morality and law has some merits and is preferable, for instance, to the mental gizmo thesis. This would be a very substantial insight for an explanatory theory of human moral cognition and an important insight for the theory of mind in general. But what *normative* importance would this have, if one wants to avoid a naturalistic fallacy? This is the last question to be asked. The answer is: It has no normative consequences as such because one needs normative arguments to justify any normative point. The mentalist theory of ethics and law refutes, however, the idea that deontology is from a hard-headed, nonarmchair, scientific point of view just a cognitive or moral illusion. Deontological arguments are not at all discredited by the theory of mind, psychology, or neuroscience. The normative task to justify any normative principle that enters into a theory of justification of human rights has not been performed by such a theory of moral cognition. It is, however, a quite different matter to defend cognitive principles of the sort outlined as reasonable, than to defend the normative value of principles that are the product of the post hoc rationalization of hard-wired emotional gut reactions. Moral psychology cannot substitute for normative theory building in ethics and law.¹⁹² But it

¹⁹² This is not only a theoretical problem. There are robust empirical findings concerning the trolley problem and the fact that most people think that turning the switch in the bystander case is permissible, saving five and foreseeing, though not directly intending that one person will be killed. In various legal systems, the application of this principle is contentious in the framework of the necessity defense. No weighing of life is permitted, e.g., for German Criminal Law, on the level of justification cf. W. Perron, in: A. Schönke/H. Schröder, *Strafgesetzbuch*, 29th ed., 2014, § 34 Rn 24. How to solve such cases is highly contentious, one solution being an extra-legal reason for exculpation, cf. T. Lenckner/D. Sternberg-Lieben, in: A. Schönke/H. Schröder, *Strafgesetzbuch*, 29th ed., 2014, Vorbemerkungen zu den 32ff, Rn 115ff; T. Hörmle, Töten, um viele Leben zu retten, in: H. Putzke (ed.), *Festschrift für Rolf Dietrich Herzberg*, 2008, p. 555ff. This was exactly the background of Welzel's invention of the trolley problem: He wanted to expand the necessity defense; cf. H. Welzel, *Zum Notstandsproblem*, ZStW 1951, p. 47ff, concretely, and that is not uninteresting (not least given his own entanglement in Nazi criminal law), for physicians involved in the killings during the Nazis' so-called euthanasia program, the mass murder of persons with disabilities. In Swiss criminal law, the situation is similar; cf. A. Coninx, *Das Solidaritätsprinzip im Lebensnotstand*, 2012, p. 55ff. Another example of real-life trolley cases is the well-known decision of the Federal German Constitutional Court striking down a regulation in the Aviation Security Act (*Luftsicherheitsgesetz*) allowing for the shooting down of an aircraft kidnapped with the

is indispensable to show that any normative theory is not just the illusionary offspring of hidden mechanisms of the mind and thus reconcilable with what is known about the structure and working of the mind.

This project of justification of human rights is a large enterprise given the complexity of the issue and the kind of considerations from anthropology and political and social theory that enter into any theoretical effort in this respect.¹⁹³ The normative justification of human rights may ultimately be based on a fallible account of certain foundational moral judgments, the reflexive resilience of which against systematic, conscious theoretical doubt is the epistemological alternative to the trilemma¹⁹⁴ of infinite regress, the dogmatic end of the justificatory argument or tautologies. The principles of justice and altruism will play an important role in this respect because there is good reason to believe that they (or some variant of them) are foundational for morality. They are not self-authenticating truths, but improvable, preliminary approximations to those principles that constitute morality.¹⁹⁵ Other principles may play an important

intention to use it as a weapon against third parties, BVerfGE 115, 118. The court argued that this is impermissible as violating the dignity of the passengers and crew of the airplane. It thus declined to apply a utilitarian calculus in this particular situation that mirrors aspects of the bystander case. Whether a restrictive conception of the necessity defense or of the reasons for exculpation or the conclusions of the court in the latter case are convincing, is not the question here. The point is that one needs normative arguments for or against these stands. Empirical theories on the structure of human moral cognition will not be sufficient (though important in the sense outlined) to answer these questions. Therefore the normative impact of theories of the mind may have to be well qualified. For some more comments with different nuances, see J. Mikhail, *Moral Grammar and Human Rights: Some Reflections on Cognitive Science and Enlightenment Rationalism*, in: R. Goodman/D. Jinks/A. K. Woods (eds.), *Understanding Social Action, Promoting Human Rights*, 2012, p. 196ff. On the problem of “normative adequacy”, cf. J. Mikhail, *Elements of Moral Cognition*, 2011, p. 29ff, 183ff.

¹⁹³ Complex and in many details historically contingent systems of rights are therefore not directly derivable from basic moral intuitions. Basic intuitions of justice or impermissible harms are one thing, the formulation and full justification of a norm in the technical form of a legal fundamental right and its regime of limitations quite another. For a derivation of a wide range of human rights from basic moral judgments, though not from principles of justice and altruism as proposed above, see J. Mikhail, *Moral Grammar and Human Rights: Some Reflections on Cognitive Science and Enlightenment Rationalism*, in: R. Goodman/D. Jinks/A. K. Woods (eds.), *Understanding Social Action, Promoting Human Rights*, 2012, p. 196ff.

¹⁹⁴ Cf. on this trilemma H. Albert, *Traktat über kritische Vernunft*, 5th ed., 1991.

¹⁹⁵ Cf. on these matters M. Mahlmann, *Ethics, Law and the Challenge of Cognitive Science*, *German Law Journal* 8 (2007), p. 593ff; idem., *Rechtsphilosophie und Rechtstheorie*, 4th ed., 2017, p. 385ff. For some comments on why one should take such judgments as foundational, cf. idem., *The Cognitive Foundations of Law-An Introduction to the Mentalist Theory of Ethics and Law*, in: H. Rottluthner, *Foundations of Law*, 2005, p. 75ff. There is the problem that certain principles of moral cognition may circumscribe the range of any valid, possible normative arguments because the latter are necessarily ultimately based on them. The reason to accept a normative proposition is, however, not that it is the product of the application of

role, too, such as the principle of noninstrumentalization of human persons or those specifying the permissibility of otherwise prohibited acts.¹⁹⁶

On this basis, further, more concrete questions can be asked: For Example, whether the ethical thought formulated in the principle of humanity and the idea of human dignity make any ethical and legal sense.¹⁹⁷ Thus, one may be able to approach a more comprehensive theory of human rights that is much needed. The argument of this chapter has been that principles of egalitarian justice, of human solidarity and care together with a sufficiently rich concept of human existence and a political theory of human flourishing embedded in a plausible theory of mind provide good reasons to believe that the idea of human rights is as well justified as anything has ever been in the history of human thought about morality and law.

Such a theory of human rights that has answers to the theoretical challenges formulated by skeptics, and is embedded in a plausible theory of the human mind, may turn out to be an important, perhaps even essential element of their intellectual defense. This would not be a small achievement. Human rights are not trivia. They are more than playthings to satisfy one's intellectual ludic drive. Human rights are not means to solve all the world's problems. But much depends on rights, including important goods of individuals, sometimes even their dignity and life. A decent level of civilization cannot be maintained without them. This is of great importance for those who suffer from human rights violations. It is of some significance as well for all those belonging to the perhaps not so small group of people who cannot breathe freely because of the continuing tragedy of human folly and pain and therefore long for the occasional relief of fresh air bestowed by some steps toward a culture of human decency.

plausible explanatory theories of moral cognition, but that the normative proposition is justified.

¹⁹⁶ The latter is the argument of J. Mikhail, *Elements of Moral Cognition*, 2011. On the common law concept of "battery" and its possible role in a mentalist ethics, idem., Any Animal Whatever? Harmful Battery and Its Elements as Building Blocks of Moral Cognition, *Ethics* 124 (2014), p. 750ff.

¹⁹⁷ On that question cf. M. Mahlmann, The Good Sense of Dignity, *Proceedings of British Academy* 192 (2013), p. 593ff.